



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 129268

TO: Michael E Szperka
Location: 3d74 / 3c70
Sunday, August 08, 2004
Art Unit: 1644
Phone: 272-2934
Serial Number: 10 / 062831

From: Jan Delaval
Location: Biotech-Chem Library
Rem 1A51
Phone: 272-2504

jan.delaval@uspto.gov

Search Notes

This Page Blank (uspto)

Delaval, Jan

129268

From: Szperka, Michael Edward
Sent: Friday, August 06, 2004 10:31 AM
To: Delaval, Jan
Subject: polypeptide search for [REDACTED] 10062831

Hi Jan!

In application 10/062831, please search:

SEQ ID NO: 59
open and closed, against the regular and interference protein databases.

[REDACTED]

Thanks!

Michael Szperka
Patent Examiner, 1644
(571) 272-2934
Remsen, 3d74

This Page Blank (uspro,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2004, 10:10:25 ; Search time 45 Seconds

(without alignments)
794.663 Million cell updates/sec

Title: US-10-062-831-59

Perfect score: 608 1 MARGSLRLRLVLGLMLA.....LSGLFLWRRCRRSSPPPX 114

Sequence: BLOSUM62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 1291235 segs, 313682936 residues

Total number of hits satisfying chosen parameters: 577798

Minimum DB seq length: 0

Maximum DB seq length: 114

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	607	99.8	114	US-10-062-831-59	Sequence 59, Appl
2	607	99.8	114	US-10-062-599-59	Sequence 139, Appl
3	274.5	45.1	112	US-10-351-334-139	Sequence 108148,
4	73.5	12.1	112	US-10-437-963-108148	Sequence 27, Appl
5	70	11.5	109	US-10-380-703-27	Sequence 61480, A
6	69.5	11.4	109	US-10-425-114-61480	Sequence 21, Appl
7	68	11.2	72	US-10-380-703-21	Sequence 163452,
8	68	11.2	72	US-10-380-703-30	Sequence 26, Appl
9	67.5	11.1	70	US-10-437-963-163452	Sequence 19, Appl
10	67	11.0	70	US-10-380-703-26	Sequence 28, Appl
11	66	10.9	70	US-10-380-703-19	Sequence 190, App
12	66	10.9	74	US-09-468-147-190	Sequence 197, App
13	66	10.9	74	US-09-468-147-197	Sequence 190, App
14	66	10.9	74	US-09-468-147-197	Sequence 190, App
15	66	10.9	74	US-10-319-745-190	Sequence 190, App

16	66	10.9	74	12	US-10-319-745-197	Sequence 197, App
17	65	10.7	70	12	US-10-380-703-22	Sequence 22, Appl
18	65	10.7	70	12	US-10-380-703-24	Sequence 24, Appl
19	65	10.7	70	12	US-10-380-703-31	Sequence 31, Appl
20	65	10.7	70	12	US-10-380-703-32	Sequence 32, Appl
21	65	10.7	87	14	US-10-029-386-28395	Sequence 28395, A
22	64	10.5	65	14	US-10-029-386-31337	Sequence 31337, A
23	63.5	10.4	114	12	US-10-424-599-153296	Sequence 153296,
24	63	10.4	70	12	US-10-380-703-20	Sequence 20, Appl
25	63	10.4	70	12	US-10-380-703-23	Sequence 23, Appl
26	63	10.4	70	12	US-10-380-703-29	Sequence 29, Appl
27	63	10.4	71	12	US-10-424-599-210769	Sequence 210769,
28	63	10.4	94	9	US-09-843-676-215	Sequence 215, App
29	63	10.4	94	10	US-09-438-486-215	Sequence 215, App
30	63	10.4	94	12	US-10-325-810-334	Sequence 334, App
31	63	10.4	94	14	US-10-053-758-215	Sequence 215, App
32	63	10.4	94	14	US-10-054-295-215	Sequence 215, App
33	63	10.4	94	14	US-10-054-611-215	Sequence 215, App
34	62.5	10.3	75	12	US-10-424-599-175619	Sequence 175619,
35	62.5	10.3	75	14	US-10-153-668-53	Sequence 53, Appl
36	62.5	10.3	112	9	US-09-864-761-36263	Sequence 36263, A
37	62.5	10.3	113	14	US-10-106-698-7944	Sequence 7944, Ap
38	62	10.2	70	12	US-10-380-703-13	Sequence 13, Appl
39	62	10.2	70	12	US-10-380-703-25	Sequence 25, Appl
40	62	10.2	101	12	US-10-424-599-152927	Sequence 152927,
41	62	10.2	103	12	US-10-424-599-217717	Sequence 217717,
42	62	10.2	114	16	US-10-437-963-132050	Sequence 132050,
43	61.5	10.1	94	16	US-10-437-963-189725	Sequence 189725,
44	61	10.0	52	12	US-10-424-599-191164	Sequence 191164,
45	61	10.0	90	14	US-10-156-761-8362	Sequence 8362, Ap

ALIGNMENTS

RESULT 1
US-10-062-831-59
; Sequence 59, Application US/10062831
; Publication No. US20030105297A1
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: P2006P1
; CURRENT APPLICATION NUMBER: US/10/062,831
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/690,454
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US98/10868
; PRIOR FILING DATE: May 28, 1998
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: SITE
LOCATION: (114)
OTHER INFORMATION: Xaa equals stop translation
US-10-062-831-59

Query Match 99.8%; Score 607; DB 14; Length 114;
Best Local Similarity 100.0%; Pred. No. 2.2e-51;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARGSLRLRLTLVGLMTALLRSVAGEQAPGTAPCSRSGSSWSADLDKCMDCASCRAPH 60
DB 1 MARGSLRLRLTLVGLMTALLRSVAGEQAPGTAPCSRSGSSWSADLDKCMDCASCRAPH 60
QY 61 SDFCLGCAAPAPPPRLIMPTILGALSLTFVLGLSGFLVWRRCRRRSSPPP 113
DB 61 SDFCLGCAAPAPPPRLIMPTILGALSLTFVLGLSGFLVWRRCRRRSSPPP 113

RESULT 2

US-10-062-599-59
Sequence 59, Application US/10062599
Publication No. US20030195346A1
GENERAL INFORMATION:
APPLICANT: Steven M. Ruben, et al.
TITLE OF INVENTION: 32 Human Secreted Proteins
FILE REFERENCE: P2006P1
CURRENT APPLICATION NUMBER: US/10/062.599
CURRENT FILING DATE: 2002-02-05
PRIOR APPLICATION NUMBER: 09/690,454
PRIOR FILING DATE: 2000-10-18
PRIOR APPLICATION NUMBER: 09/189,144
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: 60/044,039
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/048,093
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/048,190
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/050,935
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/048,101
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/048,356
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/056,250
PRIOR FILING DATE: August 29, 1997
PRIOR APPLICATION NUMBER: 60/056,296
PRIOR FILING DATE: August 29, 1997
PRIOR APPLICATION NUMBER: 60/056,293
NUMBER OF SEQ ID NOS: 229
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 59
LENGTH: 114
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (114)
OTHER INFORMATION: Xaa equals stop translation
US-10-062-599-59

Query Match 99.8%; Score 607; DB 14; Length 114;
Best Local Similarity 100.0%; Pred. No. 2.2e-51;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARGSLRLRLTLVGLMTALLRSVAGEQAPGTAPCSRSGSSWSADLDKCMDCASCRAPH 60
DB 1 MARGSLRLRLTLVGLMTALLRSVAGEQAPGTAPCSRSGSSWSADLDKCMDCASCRAPH 60
QY 61 SDFCLGCAAPAPPPRLIMPTILGALSLTFVLGLSGFLVWRRCRRRSSPPP 113

DB 61 SDFCLGCAAPAPPPRLIMPTILGALSLTFVLGLSGFLVWRRCRRRSSPPP 113

RESULT 3

US-10-351-334-139
Sequence 139, Application US/10351334
Publication No. US20040034196A1
GENERAL INFORMATION:
APPLICANT: Komatsu et al
TITLE OF INVENTION: 38 Human Secreted Proteins
FILE REFERENCE: P2031P2
CURRENT APPLICATION NUMBER: US/10/351,334
CURRENT FILING DATE: 2003-01-27
PRIOR APPLICATION NUMBER: 60/350,898
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/489,847
PRIOR FILING DATE: 2000-01-24
PRIOR APPLICATION NUMBER: PCT/US99/17130
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: 60/094,657
PRIOR FILING DATE: 1998-07-30
PRIOR APPLICATION NUMBER: 60/095,486
PRIOR FILING DATE: 1998-08-05
PRIOR APPLICATION NUMBER: 60/096,319
PRIOR FILING DATE: 1998-08-12
PRIOR APPLICATION NUMBER: 60/095,454
PRIOR FILING DATE: 1998-08-06
PRIOR APPLICATION NUMBER: 60/095,455
NUMBER OF SEQ ID NOS: 376
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 139
LENGTH: 112
TYPE: PRT
ORGANISM: Homo sapiens
US-10-351-334-139

Query Match 45.1%; Score 274.5; DB 12; Length 112;
Best Local Similarity 96.4%; Pred. No. 4.6e-19;
Matches 54; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 MARGSLRLRLTLVGLMTALLRSVAGEQAPGTAPCSRSGSSWSADLDKCMDCASCRAPH 55
DB 1 MARGSLRLRLTLVGLMTALLRSVAGEQAPGTAPCSRSGSSWSADLDKCMDCSTSC 56

RESULT 4

US-10-437-963-108148
Sequence 108148, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Mu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 108148
LENGTH: 112
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_1242C.1.pep
US-10-437-963-108148

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2004, 10:08:05 ; Search time 19 Seconds

(without alignments)
309,756 Million cell updates/sec

Title: US-10-062-831-59

608

Sequence: 1 MARGSLRLRLVLTGLMLA.....LSGFLVNRGRRRSSPPX 114

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 274203

Minimum DB seq length: 0
Maximum DB seq length: 114

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents AA:*
2: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/5B.COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6A.COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCtus.COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	607	99.8	114	4	US-09-690-454-59
2	274.5	45.1	112	4	US-09-489-847-139
3	67	11.0	112	4	US-09-252-991A-21707
4	64	10.5	95	4	US-09-252-991A-16640
5	63	10.4	94	3	US-08-851-843A-215
6	63	10.4	94	3	US-08-974-549A-334
7	63	10.4	94	3	US-08-854-050-215
8	63	10.4	94	4	US-09-430-323-215
9	63	10.4	94	4	US-09-402-181B-334
10	63	10.4	94	4	US-09-721-456-334
11	61.5	10.1	95	4	US-09-621-976-5253
12	61	10.0	71	4	US-09-107-532A-4242
13	60.5	10.0	75	6	5189019-4
14	59.5	9.8	71	4	US-09-252-991A-26803
15	59.5	9.8	71	4	US-09-252-991A-25315
16	59	9.7	68	1	US-07-696-051B-1
17	59	9.7	68	1	US-07-924-063A-1
18	59	9.7	68	1	US-08-138-340B-2
19	59	9.7	109	4	US-09-252-991A-29444
20	58.5	9.6	69	4	US-09-205-258-439
21	58.5	9.6	103	4	US-09-489-039A-7227
22	58	9.5	98	4	US-09-252-991A-31510
23	57.5	9.5	96	4	US-09-252-991A-25413
24	57.5	9.5	110	4	US-09-252-991A-18800
25	57	9.4	102	4	US-09-252-991A-22067
26	57	9.4	111	4	US-09-252-991A-17821
27	56.5	9.3	100	4	US-09-127-578-6

28	56.5	9.3	104	4	US-09-252-991A-23645	Sequence 23645, A
29	56	9.2	85	1	US-08-468-347-18	Sequence 18, Appl
30	56	9.2	85	2	US-08-467-389-18	Sequence 18, Appl
31	56	9.2	85	2	US-08-779-379-18	Sequence 18, Appl
32	56	9.2	85	2	US-08-469-219-18	Sequence 18, Appl
33	56	9.2	85	3	US-08-228-152-17	Sequence 17, Appl
34	55.5	9.1	111	4	US-09-252-991A-27861	Sequence 27861, A
35	55.5	9.1	111	4	US-09-252-991A-23578	Sequence 23578, A
36	55	9.0	79	4	US-09-621-976-6283	Sequence 6283, Ap
37	55	9.0	81	2	US-08-812-003-2	Sequence 2, Appl
38	55	9.0	92	4	US-09-127-946-8	Sequence 8, Appl
39	55	9.0	93	4	US-09-247-155-114	Sequence 114, App
40	55	9.0	93	4	US-09-148-545-203	Sequence 203, App
41	55	9.0	93	4	US-09-148-545-256	Sequence 256, App
42	55	9.0	93	4	US-09-855-323-13	Sequence 13, Appl
43	55	9.0	104	4	US-09-252-991A-31308	Sequence 31308, A
44	55	9.0	108	4	US-09-252-991A-17778	Sequence 17778, A
45	55	9.0	111	4	US-09-252-991A-23607	Sequence 23607, A

ALIGNMENTS

```
RESULT 1
US-09-690-454-59
; Sequence 59, Application US/09690454
; Patent No. 6531447
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OR INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: P2006P1
; CURRENT APPLICATION NUMBER: US/09/690,454
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patentm Ver. 2.0
; SEQ ID NO 59
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (114)
; OTHER INFORMATION: Xaa equals stop translation
US-09-690-454-59
```

```
Query Match 99.8% Score 607; DB 4; Length 114;
Best Local Similarity 100.0%; Pred. No. 3.5e-59;
Matches 113; Conservative 0; Mismatches 0; Gaps 0;
Db 1 MARGSLRLRLVLTGLMLA...LSGFLVNRGRRRSSPPX 60
```

QY 61 SDFCUGAAPPAPRLLMPLIGALSLTFVLGILSGFLVMRCRRSSPP 113
DB 61 SDFCUGAAPPAPRLLMPLIGALSLTFVLGILSGFLVMRCRRSSPP 113

RESULT 2

US-09-489-847-139
; Sequence 139, Application US/09489847
; Patent No. 64/6195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; EARLIER FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 139
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-139

Query Match 45.1%; Score 274.5; DB 4; Length 112;
Best Local Similarity 96.4%; Pred. No. 8.9e-23;
Matches 54; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 MARGSLRRLRLVLVGLMTALRSVAGEAPGTAPCSRSSWSADIDKCMDCQA-SC 55
DB 1 MARGSLRRLRLVLVGLMTALRSVAGEAPGTAPCSRSSWSADIDKCMDCSTSC 56

RESULT 3

US-09-252-991A-21707
; Sequence 21707, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21707
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21707

Query Match 11.0%; Score 67; DB 4; Length 112;
Best Local Similarity 29.2%; Pred. No. 4.8;
Matches 28; Conservative 8; Mismatches 12; Indels 48; Gaps 7;

QY 23 RSVAAGEQA---PQTAF-----CSRGSSWSADIDKCMDCASCARPHSDPCLGCA 68
DB 6 RSRPGRSRSPPCGSPATVTRGPGAGVGAR-SPW-----ASVRSRRGT-----GCG 50

QY 69 AAP-----APRLL-----WPLTGA 85
DB 51 SAPPRDPAAPGRLEFVADGRHGRGRGRGMPVAGGA 86

RESULT 4

US-09-252-991A-16640
; Sequence 16640, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16640
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16640

Query Match 10.5%; Score 64; DB 4; Length 95;
Best Local Similarity 30.0%; Pred. No. 8.4;
Matches 18; Conservative 5; Mismatches 25; Indels 12; Gaps 3;

QY 22 LRSVAGE-QAPCTARCS-----RGSWSADIDKCMDCASCARPHSDPCLGCAAPPAP 74
DB 38 LRPGGGQPRPPSAPCDPACRWRSAASWTAPAKSVACTQASVRP-----ARCAAPPGGP 92

RESULT 5

US-08-851-843A-215
; Sequence 215, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809e1 Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2004, 10:07:09 ; Search time 16 Seconds
(without alignments)
685.365 Million cell updates/sec

Title: US-10-062-831-59

Perfect score: 608

Sequence: 1 MARGSLRLRLKLVIGLMLA.....LSGLVWRRCRRSSPPPX 114

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 segs, 96191526 residues

Total number of hits satisfying chosen parameters: 47643

Minimum DB seq length: 0
Maximum DB seq length: 114

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	63	10.5	108	2 T49319	hypothetical prote
2	64	10.4	112	2 T29620	hypothetical prote
3	61	10.0	55	1 FEPE	ferredoxin 2[Fe-4
4	61	10.0	66	2 S58086	metallochionein 3
5	61	10.0	68	2 A46034	metallochionein 3
6	61	10.0	68	2 I67866	growth inhibitory
7	60.5	10.0	76	2 T03860	TA20 protein - com
8	59	9.7	78	2 B46034	metallochionein 3,
9	58.5	9.6	78	2 E69969	hypothetical prote
10	58	9.5	36	2 C45875	M1 class I histoco
11	58	9.5	36	2 S44391	metallochionein 3
12	57.5	9.5	92	2 T04395	probable phospholi
13	57	9.4	101	2 T26641	hypothetical prote
14	57	9.4	107	2 G72496	hypothetical prote
15	56.5	9.3	67	2 G82702	hypothetical prote
16	56.5	9.3	95	2 T06275	benzothiadiazole-i
17	56.5	9.3	109	2 E84202	ferredoxin [import
18	56	9.2	103	2 T08775	hypothetical prote
19	55.5	9.1	58	2 C69365	ferredoxin (fdx-5)
20	55.5	9.1	59	1 FEDV2N	ferredoxin 2[Fe-4
21	55.5	9.1	85	2 H95303	hypothetical prote
22	55.5	9.1	111	2 A85866	hypothetical prote
23	55	9.0	114	2 A13204	IS6c family orf2 f
24	54	8.9	105	2 C72608	hypothetical prote
25	54	8.9	110	2 A55991	effector cell prot
26	53.5	8.8	38	2 PS0133	H-2 class I histoc
27	53.5	8.8	81	2 S39511	photosystem I iron
28	53.5	8.8	97	2 AB2120	ferredoxin [import
29	53	8.7	59	2 I51272	retinoic acid rece

30	53	8.7	62	2 A53640	metallochionein 4
31	53	8.7	62	2 B53640	metallochionein IV
32	53	8.7	63	2 C69303	ferredoxin (fdx-4)
33	53	8.7	105	2 D82734	outer membrane pro
34	53	8.7	114	2 B60627	glycoprotein hormo
35	53	8.7	114	2 I51230	gonadotropin alpha
36	53	8.7	114	2 G89914	hypothetical prote
37	52.5	8.6	71	2 A59412	KGD-bearing platelet
38	52.5	8.6	94	2 E90451	hypothetical prote
39	52	8.6	61	2 A37425	metallochionein 2
40	52	8.6	62	2 S54335	metallochionein-2C
41	52	8.6	62	2 S54336	metallochionein-2a
42	52	8.6	65	2 F82580	hypothetical prote
43	52	8.6	81	2 A45320	transglutaminase s
44	52	8.6	86	2 E64333	ferredoxin 2[Fe-4
45	52	8.6	91	2 A59002	cryptdin-related p

ALIGNMENTS

RESULT 1
T49319
hypothetical protein B13N20.40 [imported] - Neurospora crassa
C/Species: Neurospora crassa
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C/Accession: T49319
R/Schulte, U.; Aign, V.; Hohseisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura
submitted to the Protein Sequence Database, May 2000
A/Reference number: Z25022
A/Accession: T49319
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-108 <SCH>
A/Cross-references: EMBL:AL355925; GSPDB:GN00116; NCSP:B13N20.40
A/Experimental source: BAC clone B13N20; strain OR74A
C/Genetics:
A/Gene: NCSP:B13N20.40
A/Map position: 6

Query Match 10.5%; Score 64; DB 2; Length 108;
Best Local Similarity 32.9%; Pred. No. 35;
Matches 28; Conservative 7; Mismatches 34; Indels 16; Gaps 4;

QY	13	IVLGLMLA-----LRVAGEQAPGTA-PCSRGSSWSADLDKCDKASCRARPHSDPCL	65
DB	30	LRRLRLMLTRRREMSLRMSLTPKRGGLPFSKGTITTS-----LVVLMAPPHSDCCL	82

QY 66 --GCAAPPAPPRLLMPILGALSL 88
DB 83 PTGLRARSROBSHLDPITQALDL 107

RESULT 2
T29620
hypothetical protein R10A10.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Aug-2002
A/Accession: T29620
A/Reference number: Z20653
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-112 <WAM>
A/Cross-references: EMBL:U80449; PIDN:AA837826.1; GSPDB:GN00019; CESP:R10A10.2
A/Experimental source: strain Bristol N2; clone R10A10
C/Genetics:
A/Gene: CESP:R10A10.2
A/Map position: 1
A/Introns: 17/2, 59/1

C:Superfamily: yeast hypothetical protein YOL133w; RING finger homology

Query Match 10.4%; Score 63; DB 2; Length 112;
Best Local Similarity 27.5%; Pred. No. 45;
Matches 19; Conservative 8; Mismatches 22; Indels 20; Gaps 3;

QY 11 RLIVLGLMLTSLVAGEQAPGAPRCRSGSSWADLCKMDCASCRARPHSDPCLGCAAA 70
DB 28 RPPVLKMKNNL-----AYMNDVE-CDTCAICRVHLMEE-CLRQSE 67

QY 71 PPAPFRLW 79
DB 68 PSABCYVM 76

RESULT 3

PEEP
ferredoxin 2[4Fe-4S] [validated] - Peptostreptococcus asaccharolyticus

C:Species: Peptostreptococcus asaccharolyticus

C>Date: 24-Apr-1984 #sequence_revision 23-Mar-1995 #text_change 20-Apr-2000

C:Accession: A00196

R:Funoda, J.N.; Yasunobu, K.T.; Whiteley, H.R.

J. Biol. Chem. 243, 6262-6272, 1968

A:Title: Non-heme iron proteins. IX. The amino acid sequence of ferredoxin from Micrococcus

A:Reference number: A92040; MUID:59054261; PMID:5723466

A>Note: the source is designated as Micrococcus aerogenes

A:Accession: A00196

A:Molecule type: protein

A:Residues: 1-21,23-24, 'O', 26-55 <TSU>

R:Backes, G.; Mino, Y.; Loefer, T.M.; Meyer, T.E.; Cusanovich, M.A.; Sweeney, W.V.; Adman

J. Am. Chem. Soc. 113, 2055-2064, 1991

A:Title: The environment of FeS4 clusters in ferredoxins and high-potential iron protei

A:Reference number: A44688

A:Contents: annotation; X-ray crystallography; 2.0 angstroms; sequence revision

A>Note: sequence correction confirmed by peptide sequencing

R:Adman, E.T.; Sieker, L.C.; Jensen, L.H.

submitted to the Brookhaven Protein Data Bank, August 1976

A:Reference number: A50836; PDB:1FDX

A:Contents: annotation; X-ray crystallography; 2.0 angstroms; residues 1-21, 'I', 23-24, 26

R:Adman, E.T.; Sieker, L.C.; Jensen, L.H.

J. Biol. Chem. 251, 3801-3806, 1976

A:Title: Structure of Peptococcus aerogenes ferredoxin. Refinement at 2 angstroms resolu

A:Reference number: A92192; MUID:76213238; PMID:932007

A:Contents: annotation; X-ray crystallography; 2.0 angstroms

R:Adman, E.T.; Sieker, L.C.; Jensen, L.H.

J. Biol. Chem. 248, 3987-3996, 1973

A:Title: The structure of a bacterial ferredoxin.

A:Reference number: A92136; MUID:73187389; PMID:4708097

A:Contents: annotation; X-ray crystallography; 2.8 angstroms

C:Superfamily: ferredoxin 2[4Fe-4S]; ferredoxin 2[4Fe-4S] homology

C:Keywords: 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein

P:1-54/Domain: ferredoxin 2[4Fe-4S] homology <PBR>

F:8,11,14,46/Binding site: 4Fe-4S cluster (Cys) (covalent) #status experimental

F:18,36,39,42/Binding site: 4Fe-4S cluster (Cys) (covalent) #status experimental

Query Match 10.0%; Score 61; DB 1; Length 55;
Best Local Similarity 33.3%; Pred. No. 41;
Matches 12; Conservative 4; Mismatches 12; Indels 8; Gaps 1;

QY 39 GSKMSADLCKMDCASCRARPHSDPCLGCAAPAP 74
DB 26 GSIVADIDSCIDGSCAS-----VCPVGAAP 53

RESULT 4
S58086
metallothionein 3 - rat

N:Alternate names: neurotrophic growth inhibitory factor

C:Species: Rattus norvegicus (Norway rat)

C>Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 20-Aug-1999

C:Accession: S58086; I52636

R:Amoureux, M.C.; Rehsaus, E.; Wurich, T.; Colpaert, F.C.; Pauwels, P.J.

submitted to the EMBL Data Library, July 1995

A:Reference number: S58084

A:Accession: S58086

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-66 <ANO>

A:Cross-references: EMBL:X89603; NID:9908880; PIDN:CA61762.1; PID:9908881

R:Kobayashi, H.; Uchida, Y.; Ihara, Y.; Nakajima, K.; Kohsaka, S.; Miyatake, T.; Tsuji,

Brian Res. Mol. Brain Res. 19, 188-194, 1993

A:Title: Molecular cloning of rat growth inhibitory factor cDNA and the expression in t

A:Reference number: I52636; MUID:94018480; PMID:8412560

A:Accession: I52636

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-66 <KOB>

A:Cross-references: GB:S65838; NID:9425381; PIDN:AAB28366.1; PID:9425382

C:Superfamily: metallothionein

C:Keywords: acetylated amino end; chelation; metal binding; metal-thiolate cluster

F:1/Modified site: acetylated amino end (Met) #status predicted

F:6,8,14,16,20,22,25,27,30/Binding site: transition metal ions (Cys) #status predicted

F:34,35,37,38,42,45,49,51,62,64,65/Binding site: transition metal ions (Cys) #status pr

Query Match 10.0%; Score 61; DB 2; Length 66;
Best Local Similarity 35.0%; Pred. No. 47;
Matches 14; Conservative 2; Mismatches 22; Indels 2; Gaps 2;

QY 31 PGTAPCRSGSSWADLCKMDCASCRARPHSDPCLGCAAA 70
DB 3 PFTCPPTGSGCTCS-DKC-KCKGCKTCKCKSCSCCPA 40

RESULT 5
A46034
metallothionein 3, brain-specific - mouse

N:Alternate names: neurotrophic growth inhibitory factor

C:Species: Mus musculus (house mouse)

C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999

C:Accession: A46034

R:Palmiter, R.D.; Findley, S.D.; Whitmore, T.E.; Durnam, D.M.

Proc. Natl. Acad. Sci. U.S.A. 89, 6333-6337, 1992

A:Title: MT-III, a brain-specific member of the metallothionein gene family.

A:Reference number: A46034; MUID:92335292; PMID:1631128

A:Accession: A46034

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-68 <PAL>

A:Cross-references: GB:M93310; NID:9199133; PIDN:AAA3529.1; PID:9199134

A>Note: sequence extracted from NCBI backbone (NCBIN:108715, NCBIN:111115, NCBIN:108716)

C:Superfamily: metallothionein

Query Match 10.0%; Score 61; DB 2; Length 68;
Best Local Similarity 35.0%; Pred. No. 48;
Matches 14; Conservative 2; Mismatches 22; Indels 2; Gaps 2;

QY 31 PGTAPCRSGSSWADLCKMDCASCRARPHSDPCLGCAAA 70
DB 3 PFTCPPTGSGCTCS-DKC-KCKGCKTCKCKSCSCCPA 40

RESULT 6
I67866
growth inhibitory factor - mouse

C:Species: Mus sp. (mouse)

C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 20-Aug-1999

C:Accession: I67866

R:Maruse, S.; Igataashi, S.; Furuya, T.; Kobayashi, H.; Miyatake, T.; Tsuji, S.

Gene 144, 283-287, 1994

A:Title: Structures of the human and mouse growth inhibitory factor-encoding genes.

A:Reference number: I53803; MUID:94314230; PMID:8039715

A:Accession: I67866

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-68 <RBS>

A:Cross-references: GB:S72046; NID:9565191; PIDN:AAB31397.1; PID:9565192

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2004, 10:03:29 ; Search time 13 Seconds

(without alignments)
456.615 Million cell updates/sec

Title: US-10-062-831-59

Sequence: 1 MARGSLRLRLVGLMIA.....LSGLVWRRCRRSSPPPX 114

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 18957

Minimum DB seq length: 0
Maximum DB seq length: 114

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73.5	12.1	81	1	PORD METHH
2	64.5	10.6	112	1	PLA8 MOUSE
3	62	10.2	54	1	FER PERAS
4	61	10.0	66	1	MT3 FAT
5	61	10.0	68	1	MT3 MOUSE
6	59.5	9.8	100	1	TXOB ATRIL
7	59	9.7	68	1	MT3 HUMAN
8	58.5	9.6	78	1	YOZF BACSV
9	58	9.5	68	1	MT3 BOVIN
10	57	9.4	100	1	RIS SHEEP
11	57	9.4	108	1	NIFZ RHISN
12	56.5	9.3	100	1	TXOA ATRIL
13	56	9.2	95	1	GONI PAGWA
14	56	9.2	113	1	RX2 MOUSE
15	55.5	9.1	59	1	FER2 DESDN
16	55	9.0	113	1	RX2 HUMAN
17	53.5	8.8	80	1	PSAC ANISP
18	53.5	8.8	97	1	FER3 ANASP
19	53.5	8.8	111	1	YFBW ECOI
20	53	8.7	62	1	MT4 CANFA
21	53	8.7	62	1	MT4 HUMAN
22	53	8.7	62	1	MT4 MOUSE
23	53	8.7	86	1	COX3 ANAPL
24	52	8.7	114	1	GLH2 ONCKE
25	52	8.6	62	1	MT2A RABIT
26	52	8.6	62	1	MT2C RABIT
27	52	8.6	81	1	TRGS TACTR
28	52	8.6	86	1	PORD METHA
29	52	8.6	91	1	DEFY MOUSE
30	52	8.6	98	1	PORD THEMVA
31	51.5	8.6	104	1	PORD CYRPU
32	51.5	8.5	81	1	PSAC CYAME
33	51.5	8.5	100	1	S115_HUMAN

34	51.5	8.5	111	1	YFBW SALTI	P81891 salmonella
35	51.5	8.5	111	1	YFBW SALTY	O52328 salmonella
36	51	8.4	55	1	FER GLOSP	P00197 clostridium
37	51	8.4	55	1	FER_CLOPS	P00200 clostridium
38	51	8.4	58	1	MT1 SCYSE	P02805 scylla serr
39	51	8.4	59	1	FER_ENTHI	P11425 entamoeba h
40	51	8.4	59	1	MT1 CALSI	P55949 callinectes
41	51	8.4	68	1	MT3 HORSE	P37360 equus caball
42	51	8.4	68	1	MT3_PIG	P55944 sus scrofa
43	51	8.4	75	1	SCS8_DROME	P02842 drosophila
44	51	8.4	83	1	MTX1_DROME	P23487 drosophila
45	51	8.4	87	1	Y476_MYCLE	Q9c4b3 mycobacteri

ALIGNMENTS

RESULT 1	ID	PORD METHH	STANDARD;	PRT;	81 AA.
AC	P56815;				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Pyruvate synthase subunit pord (BC 1.2.7.1) (Pyruvate oxidoreductase delta chain) (POR) (Pyruvic-ferredoxin oxidoreductase delta subunit).				
GN	PORD OR MTH1740.1.				
OS	Methanobacterium thermoautotrophicum.				
OC	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;				
OC	Methanobacteriaceae; Methanothermobacter.				
ON	NCBI_TaxID=187420;				
RP	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Delta H;				
RX	MEDLINE=98037514; PubMed=9371463;				
RA	Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,				
RA	Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,				
RA	Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,				
RA	Spadafora R., Viccare R., Wang Y., Wierzbowski J., Gibson R.,				
RA	Jiwani N., Carnes A., Bush D., Safer H., Patwell D., Prabhakar S.,				
RA	McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,				
RA	Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;				
RT	"Complete genome sequence of Methanobacterium thermoautotrophicum				
RT	delta: functional analysis and comparative genomics.";				
RL	J. Bacteriol. 179:7135-7155(1997).				
CC	-1- CATALYTIC ACTIVITY: Pyruvate + CoA + oxidized ferredoxin = acetyl-				
CC	CoA + CO(2) + reduced ferredoxin.				
CC	-1- COFACTOR: Binds 2 4Fe-4S clusters.				
CC	-1- SUBUNIT: Heterotetramer of one alpha, one beta, one delta and one				
CC	gamma chain.				
CC	-1- SIMILARITY: Belongs to the bacterial-type ferredoxin family.				
CC	-1- CAUTION: THERE SEEMS TO BE A SEQUENCING ERROR THAT FUSES TOGETHER				
CC	PORD AND PORD. WE HAVE CUT THE ORF INTO ITS TWO CONSTITUENTS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by, and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
CC	EMBL; AEO00929; AAB86210.1; ALT_INTR.				
DR	HSSP; P00195; ICLP.				
DR	InterPro: IPR001450; 4Fe4S_ferredoxin.				
DR	Pfam; PF00037; fer4; 2.				
DR	PROSITE; PS00198; 4Fe4S_FERREDOXIN; 2.				
KW	Oxidoreductase; Electron transport; Iron-sulfur; Repeat; 4Fe-4S;				
KW	Complete proteome.				
FT	METAL	34			IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT	METAL	37			IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT	METAL	40			IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT	METAL	44			IRON-SULFUR 2 (4FE-4S) (POTENTIAL).

```

FT METAL 60 60 IRON-SULFUR 2 (4FE-4S) (POTENTIAL)
FT METAL 63 63 IRON-SULFUR 2 (4FE-4S) (POTENTIAL)
FT METAL 66 66 IRON-SULFUR 2 (4FE-4S) (POTENTIAL)
FT METAL 70 70 IRON-SULFUR 1 (4FE-4S) (POTENTIAL)
SQ SEQUENCE 81 AA; 9121 MW; 219A9CACA8A41604 CRC64;

Query Match 12.1%; Score 73.5; DB 1; Length 81;
Best Local Similarity 29.2%; Pred. No. 1.2;
Matches 19; Conservative 8; Mismatches 21; Indels 17; Gaps 4;

QY 31 PGTAFCSSGSSWSA-----DLKCMDCASCRA-----RPHS---DFGLGCA-AAAPA 73
DB 11 PGSTYANKGSKWRTKRPVLDKXKCIDCNCLFCEGGINRHHHDYDKCKGCGIACKXC 70
QY 74 PERLL 78
DB 71 PVKAI 75

RESULT 2
PLAB MOUSE STANDARD; PRT; 112 AA.
AC Q9J148;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Placenta-specific gene 8 protein (C15 protein) (Onzin).
GN Plac8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherwin J.R.A., Sharkey A.M., Smith S.K.;
RT Identification of LIF regulated genes in the mouse uterus.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=22271112; PubMed=12384430;
RA Rissom M.-C., Duhon T., Bridon J.-M., Bendriss-Vermare N.,
RA Perome C., de Saint Vis B., Briere F., Bates E.B.M.;
RT "Subtractive hybridization reveals the expression of
RT immunoglobulinlike transcript 7, Bph-B1, granzyme B, and 3 novel
RT transcripts in human plasmacytoid dendritic cells.";
RL Blood 100:3295-3303(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=pancreas;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawal U., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Koehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Sakurai L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
RA Blake J., Boffelli D., Bojunga N., Carrincci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Rodriguez I., Sakamoto N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;

```

```

RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner U., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi F., Caavaant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carrincci P., Prange C.,
RA Rana S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1 SIMILARITY: Belongs to the UPF0281 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF263458; AAF6887.1; -
DR EMBL; AJ422151; CAD19534.1; -
DR EMBL; AK007363; BMB24991.1; -
DR EMBL; BC010789; AAH10789.1; -
DR MGI; MGI:2445289; Plac8.
DR InterPro; IPR006461; DUF_A_thal_Cys.
DR Pfam; PF04749; DUF614; 1
DR TIGRfam; TIGR01571; A_thal_Cys-rich; 1.
SQ SEQUENCE 112 AA; 12363 MW; 4FFB558BD1F0DA8 CRC64;

Query Match 10.6%; Score 64.5; DB 1; Length 112;
Best Local Similarity 34.1%; Pred. No. 12;
Matches 14; Conservative 2; Mismatches 24; Indels 1; Gaps 1;

QY 31 PGTAFCSSGSSWSADLDCM-DCASGRARHSDFCGCAA 70
DB 12 PGFVRAPQNSNMWTSLCDPDSOGVCLGTFPTCLGCGVA 52

RESULT 3
PER PEAPS STANDARD; PRT; 54 AA.
AC P00193;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ferredoxin.
OS Peptostreptococcus asaccharolyticus (Peptococcus asaccharolyticus).
OC Bacteria; Firmicutes; Clostridia; Clostridiales;
OC Peptostreptococcaceae; Peptoniphilus.
OX NCBI_Taxid=1258;
RN [1]
RP SEQUENCE.
RA MEDLINE=69054261; PubMed=5723466;
RA Tsunoda U.N., Yasunobu K.T., Whiteley H.R.;
RT "Non-heme iron proteins. IX. The amino acid sequence of ferredoxin
RT from Micrococcus aerogenes.";
RL J. Biol. Chem. 243:6262-6272(1968).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=73187389; PubMed=4708097;
RA Adman E.T., Sliker L.C., Jensen L.H.;

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2004, 10:06:34 ; Search time 37 Seconds

(without alignments)
972.137 Million cell updates/sec

Title: US-10-062-831-59

Perfect score: 608
Sequence: 1 MARGSRRLRLVLVGLMLA.....LSGLVWRRCRRSSPPPX 114

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 213498

Minimum DB seq length: 0
Maximum DB seq length: 114

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	11.0	83	17	Q8ZTF5 Pyrobaculum
2	67	11.0	90	16	Q7U455 Pyrobaculum
3	65	10.7	105	16	Q884Q0 Pseudomonas
4	65	10.7	107	12	Q36606 Pseudomonas
5	65	10.7	107	12	Q36605 Pseudomonas
6	64	10.5	77	16	Q92MCI Rhizobium m
7	63.5	10.4	85	1	Q9P9B6 Pseudomonas
8	63	10.4	112	5	P91404 Pseudomonas
9	61	10.0	90	16	Q82PP8 Streptomyces
10	60.5	10.0	76	10	Q24155 Nicotiana glauca
11	60.5	10.0	89	17	Q8T2S8 Methanopyrus
12	60.5	10.0	101	16	Q9AXI6 Oryza sativa
13	60	9.9	100	16	Q7VBB2 Prochlorococcus
14	59.5	9.8	112	10	Q7XSY1 Oryza sativa
15	59	9.7	59	12	Q9W7U2 Hepatitis e
16	59	9.7	94	2	Q8VUF9 Azotobacter ev

17	58.5	9.6	58	12	Q8UTL4 Hepatitis e
18	58.5	9.6	95	10	Q84Q89 Trilicium ae
19	58.5	9.6	107	17	Q8TYE4 Oryza sativa
20	58.5	9.6	109	10	Q8W5F8 Oryza sativa
21	58	9.5	104	10	Q8GVT9 Oryza sativa
22	58	9.5	105	16	Q7UNC9 Rhodospirillum rubrum
23	57.5	9.5	92	10	Q81135 Hordeum vulgare
24	57	9.4	69	6	Q28101 Bos taurus
25	57	9.4	77	4	Q96D32 Homo sapiens
26	57	9.4	101	5	Q9XWU9 Caenorhabditis elegans
27	57	9.4	101	15	Q8UTD0 Homo sapiens
28	57	9.4	106	17	Q69368 Streptomyces
29	57	9.4	107	12	Q9Y8L9 Streptomyces
30	57	9.4	113	2	Q83VB6 Streptomyces
31	56.5	9.3	67	16	Q9PDU8 Xylella fastidiosa
32	56.5	9.3	86	16	Q82NH5 Streptomyces
33	56.5	9.3	95	10	Q41521 Trilicium ae
34	56.5	9.3	97	4	Q86SR0 Homo sapiens
35	56.5	9.3	109	17	Q9HS23 Halobacterium salinarum
36	56	9.2	72	15	Q75047 Homo sapiens
37	56	9.2	96	9	Q8LTD0 Homo sapiens
38	56	9.2	101	15	Q8UT22 Homo sapiens
39	56	9.2	103	4	Q9Y3X7 Homo sapiens
40	55.5	9.1	33	4	Q9UD12 Homo sapiens
41	55.5	9.1	44	12	Q9Y1J3 Swine hepatitis virus
42	55.5	9.1	58	17	Q29339 Archaeoglobus fulgidus
43	55.5	9.1	85	16	Q922W2 Rhizobium m
44	55.5	9.1	108	16	Q8NMW8 Corynebacterium jeikeium
45	55.5	9.1	110	12	Q9IR98 Japanese rice

ALIGNMENTS

RESULT 1

Q8ZTF5 PRELIMINARY; PRT; 83 AA.
AC Q8ZTF5;
DT 01-MAR-2002 (TREMBL) 20, Created
DT 01-MAR-2002 (TREMBL) 20, Last sequence update
DT 01-JUN-2003 (TREMBL) 24, Last annotation update
DE 2-ketolactone ferredoxin oxidoreductase delta subunit.
GN PAE3276.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;
RA Fite-Gibson S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I., Miller J.H.,
"Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum aerophilum.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
RL EMBL; AB009922; AAL64806.1; -
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001450; 4Fe4S_ferredoxin.
DR Pfam; PF00037; fer4_2.
DR PROSITE; PS00198; 4Fe4S_FERREDOXIN; 2.
KW Complete proteome.
SQ SEQUENCE 83 AA; 8999 MW; 11E95466C4807A37 CRC64;

Query Match 11.0%; Score 67; DB 17; Length 83;
Best Local Similarity 30.4%; Pred. No. 16;

Matches 17; Conservative 4; Mismatches 13; Indels 22; Gaps 3;

34 APCSRG--SSWS-----ADLDKMDKASC-----RAPHSPFCGC 67
Db 8 APASAGITGWTYKPVVNLKIDGLCWLYCPESVIMDKTKVRIIDYDFCKGC 63

[illegible]

```

Db      24 IGLIALLLAVSQA-----LLTVAGVSSLVGHPC---MLGELINASP-- 61
QY      75 FRLIMPLIGALSLTFVLGLISGLF--VMRRCRESS 110
Db      62 -----GGA-AFVAVSGLLVFLYGLVIMRRCRORSA 91

RESULT 4
036606 ID 036606 PRELIMINARY; PRT; 107 AA.
AC 036606;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Strain Morocco/F12 ORF-1 (Fragment).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Morocco/F12;
RX MEDLINE=97475412; PubMed=9334924;
RA Chatterjee R., Tsarev S., Pillot J., Coursaget P., Emerson S.U.,
RA Purcell R.H.;
RT "African strains of hepatitis E virus that are distinct from Asian
RT strains.";
RL J. Med. Virol. 53:139-144(1997).
DR EMBL; AF010429; AAB6553.1; -.
FT NON TER 1 1
FT NON TER 107 107
SQ SEQUENCE 107 AA; 10934 MW; 7F60EDCE2E56E7F8 CRC64;

Query Match 10.7%; Score 65; DB 12; Length 107;
Best Local Similarity 32.8%; Pred. No. 32;
Matches 21; Conservative 4; Mismatches 23; Indels 16; Gaps 2;

QY      83 GGAL 86
Db      76 AAPL 79

RESULT 5
036605 ID 036605 PRELIMINARY; PRT; 107 AA.
AC 036605;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Strain Morocco/F23 ORF-1 (Fragment).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Morocco/F23;
RX MEDLINE=97475412; PubMed=9334924;
RA Chatterjee R., Tsarev S., Pillot J., Coursaget P., Emerson S.U.,
RA Purcell R.H.;
RT "African strains of hepatitis E virus that are distinct from Asian
RT strains.";
RL J. Med. Virol. 53:139-144(1997).
DR EMBL; AF010428; AAB6552.1; -.
FT NON TER 1 1
FT NON TER 107 107
SQ SEQUENCE 107 AA; 10934 MW; 7F60EDCE2E56E7F8 CRC64;

Query Match 10.7%; Score 65; DB 12; Length 107;

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2004, 10:02:59 ; Search time 52 Seconds

(without alignments)
619.431 Million cell updates/sec

Title: US-10-062-831-59

Perfect score: 608
Sequence: 1 MARGSLRLRLVLGLMLA.....LSGLVWRRCRRKRSPPPX 114

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 931724

Minimum DB seq length: 0
Maximum DB seq length: 114

Post-processing: Minimum March 0%
Maximum March 100%
Listing first 45 summaries

Database :

1: Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	607	99.8	114	2	AAW73409 Human sec
2	274.5	45.1	112	3	AAy91463 Human sec
3	274.5	45.1	112	6	ADAS7390 Human sec
4	274.5	45.1	112	7	ADCT4462 Human sec
5	274.5	45.1	112	7	ADD38025 Human sec
6	191	31.4	32	6	ADDA49370 Human sec
7	73.5	12.1	98	6	AAU66196 Human Fnl
8	73.5	12.1	98	6	ABM62715 Propionib
9	70	11.5	79	5	ABE22260 Human BAF
10	69	11.3	79	5	ABP06392 Human ORF
11	68.5	11.3	101	2	AAU57190 Human pro
12	68.5	11.3	114	4	AAU57190 Human pro
13	68.5	11.3	114	6	ABM53709 Propionib
14	68	11.2	70	5	AAE22263 Propionib
15	68	11.2	70	5	AAE22263 Propionib
16	67.5	11.1	74	4	AAU47590 Human BAF
17	67.5	11.1	74	4	AAU47590 Human BAF
18	67	11.0	70	5	AAE22259 Propionib
19	67	11.0	111	4	AAU42650 Human BAF
20	67	11.0	111	6	ABM39169 Propionib
21	66.5	10.9	98	4	AAU47442 Propionib
22	66.5	10.9	98	6	ABM43961 Propionib
23	66	10.9	70	5	AAE22261 Human BAF
24	66	10.9	70	5	AAE22252 Human BAF
25	66	10.9	74	2	AAy31399 HEV pepti

26	66	10.9	74	2	AAy31398 HEV pepti
27	65	10.7	70	5	AAE22255 Human BAF
28	65	10.7	70	5	AAE22257 Human BAF
29	65	10.7	70	5	AAE22265 Human BAF
30	65	10.7	70	5	AAE22265 Human BAF
31	64.5	10.6	61	4	AAE22264 Human BAF
32	64.5	10.6	61	6	AAU53825 Propionib
33	64	10.5	77	2	AAU53825 Propionib
34	64	10.5	100	4	AAU53825 Propionib
35	63.5	10.4	78	4	AAU53825 Propionib
36	63.5	10.4	78	6	AAU53825 Propionib
37	63.5	10.4	85	4	AAU53825 Propionib
38	63	10.4	70	5	AAU53825 Propionib
39	63	10.4	70	5	AAU53825 Propionib
40	63	10.4	70	5	AAU53825 Propionib
41	62.5	10.3	95	4	AAU53825 Propionib
42	62.5	10.3	95	6	AAU53825 Propionib
43	62.5	10.3	96	6	AAU53825 Propionib
44	62.5	10.3	96	6	AAU53825 Propionib
45	62.5	10.3	112	4	AAU53825 Propionib

ALIGNMENTS

RESULT 1	AAW73409	standard; protein; 114 AA.
ID	AAW73409	
XX	AAW73409;	
AC	AAW73409;	
DT	19-FEB-1999	(first entry)
XX		
DE	Human secreted protein encoded by Gene No. 13.	
XX		
KM	Secreted protein; human; protein therapy; blood disorder;	
KW	pathological condition; diagnosis; cancer; neurological disorder;	
KW	developmental abnormality; foetal deficiency; leukaemia; hepatic disease;	
KW	immune system disorder; Alzheimer's disease; cognitive disorder;	
KW	schizophrenia; prostate disease; autoimmune disorder; AIDS.	
XX		
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	Misc-difference 114	/note= "unspecified amino acid"
XX		
PN	W09854206-A1.	
XX		
XX	03-DEC-1998.	
PD		
XX		
PF	28-MAY-1998;	98WC-US010868.
XX		
PR	30-MAY-1997;	97US-0044039P.
PR	30-MAY-1997;	97US-0048039P.
PR	30-MAY-1997;	97US-0048101P.
PR	30-MAY-1997;	97US-0048190P.
PR	30-MAY-1997;	97US-0048356P.
PR	30-MAY-1997;	97US-0050335P.
PR	28-AUG-1997;	97US-0056250P.
PR	29-AUG-1997;	97US-0056293P.
PR	29-AUG-1997;	97US-0056296P.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Ruben SM, Rosen CA, Carter KC, Dillon PJ, Endress GA, Yu G;	
PI	Ni J, Feng P;	
XX		
DR	WPI; 1999-070209/06.	
XX		
XX	N-PSDB; AAV08823.	
PT	New isolated human genes - useful for diagnosis and treatment of, e.g.	
PT	cancers, neurological disorders, immune diseases, developmental disorders	

PT or blood disorders.
 XX
 PS Claim 11; Page 153; 188pp; English.
 CC This sequence is encoded by a cDNA of the invention, designated Gene No.
 CC 13. This sequence represents a human secreted protein, and is expressed
 CC in keratinocytes and to a lesser extent in endothelial cells and
 CC placenta. The DNA sequences of the invention and their corresponding
 CC secreted polypeptides are useful for preventing, treating or ameliorating
 CC medical conditions, e.g. by protein or gene therapy. Also pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the DNA sequences. Specific uses are described for each of the DNA
 CC sequences and the encoded proteins, based on which tissues they are most
 CC highly expressed in, and include developing products for the diagnosis or
 CC treatment of cancer, tumors, neurological disorders, developmental
 CC abnormalities and foetal deficiencies, blood disorders, leukemias,
 CC diseases of the immune system (including allergies or asthma), hepatic
 CC disease, Alzheimer's and cognitive disorders, schizophrenia, prostate
 CC diseases, autoimmune disorders and AIDS. The polypeptides are also useful
 CC for identifying their binding partners
 CC
 SQ Sequence 114 AA:
 Query Match 99.8%; Score 607; DB 2; Length 114;
 Best Local Similarity 100.0%; Pred. No. 1.3e-55; Mismatches 0; Gaps 0;
 Matches 113; Conservative 0; Indels 0;
 QY 1 MARGSLRLRLVLVGLMTALRSVAGEAPGTAPCSRSGSSWSADIDKCMDCASCRAPH 60
 DB 1 MARGSLRLRLVLVGLMTALRSVAGEAPGTAPCSRSGSSWSADIDKCMDCASCRAPH 60
 QY 61 SDFCTGCAAPAPPRLLMPTITGALSTFTVGLISGFIWRCRCRRERSPP 113
 DB 61 SDFCTGCAAPAPPRLLMPTITGALSTFTVGLISGFIWRCRCRRERSPP 113
 RESULT 2
 AA91463
 ID AA91463 standard; protein; 112 AA.
 AC
 XX AA91463;
 XX
 DT 29-JUN-2000 (first entry)
 XX
 DE Human secreted protein sequence encoded by gene 13 SEQ ID NO:136.
 XX
 KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
 KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
 KW antipsoriatic; cardiac; gene therapy; cancer; neurological disorder;
 KW immune disease; inflammation; blood disorder; tumour; chromosome 16.
 XX
 OS Homo sapiens.
 XX
 PN WO200006698-A1.
 PD 10-FEB-2000.
 XX
 PF 29-JUL-1999; 99WO-US017130.
 XX
 PR 30-JUL-1998; 98US-0094657P.
 PR 05-AUG-1998; 98US-0095486P.
 PR 06-AUG-1998; 98US-0095454P.
 PR 06-AUG-1998; 98US-0095455P.
 PR 12-AUG-1998; 98US-0096319P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
 PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
 PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
 XX

DR WPI: 2000-195282/17.
 DR N-PSDB; AAA26358.
 XX
 PT New isolated human genes and the secreted polypeptides they encode,
 XX useful for diagnosis and treatment of e.g. cancer, neurological
 XX disorders, immune diseases, inflammation or blood disorders.
 PS Claim 11; Page 465; 634pp; English.
 CC The polynucleotide sequences given in AAA26346 to AAA26458 encode the
 CC human secreted proteins given in AA91451 to AA91461. The human secreted
 CC proteins can have activities based on the tissues and cells they are
 CC expressed in. Examples of the activities are: cytostatic;
 CC immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
 CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
 CC antiasthma; antipsoriatic; and cardiac. The polynucleotides and their
 CC corresponding secreted proteins are useful for preventing, treating or
 CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
 CC pathological conditions can be diagnosed by determining the amount of the
 CC proteins in a sample or by determining the presence of mutations in the
 CC polynucleotides. Specific uses are described for each of the
 CC polynucleotides, based on which tissues they are most highly expressed
 CC in, and include developing products for the diagnosis or treatment of
 CC cancer, tumors, neurodegenerative disorders, developmental abnormalities
 CC and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, inflammation, allergies,
 CC Alzheimer's and behavioural disorders, schizophrenia, osteoporosis,
 CC arthritis, infections, AIDS, spinal cord injuries, transplant rejection,
 CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
 CC reproductive disorders, gastrointestinal disorders, respiratory disorders
 CC and metabolic disorders. The proteins or polynucleotides can also be used
 CC as food additives or preservatives. The proteins are also useful for
 CC identifying their binding partners. AAA26337 to AAA26345 and AA91450 are
 CC sequences used in the exemplification of the present invention
 CC
 SQ Sequence 112 AA:
 Query Match 45.1%; Score 274.5; DB 3; Length 112;
 Best Local Similarity 96.4%; Pred. No. 7.2e-21; Mismatches 54; Conservative 1; Indels 1; Gaps 1;
 QY 1 MARGSLRLRLVLVGLMTALRSVAGEAPGTAPCSRSGSSWSADIDKCMDCASCRAPH 60
 DB 1 MARGSLRLRLVLVGLMTALRSVAGEAPGTAPCSRSGSSWSADIDKCMDCASCRAPH 60
 RESULT 3
 ADA57390
 ID ADA57390 standard; protein; 112 AA.
 AC
 XX ADA57390;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human secreted protein #172.
 XX
 KW immunosuppressive; antiinflammatory; antiasthmatic; antiallergic;
 KW cytostatic; cerebroprotective; neuroprotective; nootropic;
 KW cardiovascular; antiarteriosclerotic; gene therapy;
 KW human secreted protein; immune disorder; inflammation;
 KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
 KW inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
 KW multiple sclerosis; ischemic brain injury; Parkinson's disease;
 KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
 KW triple helix formation; antisense gene therapy; forensic biology.
 XX
 OS Homo sapiens.
 XX
 PN WO2002102994-A2.
 PD 27-DEC-2002.
 XX
 PF 19-MAR-2002; 2002WO-US008278.
 XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2004, 10:02:34 ; Search time 46 Seconds

(without alignments)
777.388 Million cell updates/sec

Title: US-10-062-831-59

Sequence: 1 MARGSLRLRLRLVGLMLA.....LSGLVWRKRKRSPPPX 114

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1291235 segs, 313682936 residues

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10C_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	607	99.8	114	14	US-10-062-831-59
2	607	99.8	114	14	US-10-062-831-59
3	571	93.9	129	9	US-09-742-454A-4
4	571	93.9	129	9	US-09-742-454A-4
5	571	93.9	129	14	US-10-024-298A-178
6	571	93.9	129	14	US-10-042-211A-178
7	571	93.9	129	15	US-10-331-496A-37
8	571	93.9	129	15	US-10-295-027-444
9	571	93.9	129	15	US-10-295-027-1305
10	571	93.9	129	16	US-10-617-217A-178
11	450	74.0	129	9	US-09-742-454A-5
12	450	74.0	129	9	US-09-883-777-5
13	433	71.2	309	9	US-09-742-454A-7
14	433	71.2	309	9	US-09-883-777-7
15	379.5	62.4	300	9	US-09-883-777-9

16	274.5	45.1	112	12	US-10-351-334-139
17	274.5	45.1	155	12	US-10-351-334-284
18	274.5	45.1	156	12	US-10-351-334-228
19	97.5	16.0	171	14	US-10-251-947-4
20	97.5	16.0	171	14	US-10-251-947-7
21	97.5	16.0	185	14	US-10-251-947-2
22	97.5	16.0	170	14	US-10-251-947-6
23	96.5	15.9	186	14	US-10-251-947-14
24	88.5	14.6	185	12	US-10-380-703-5
25	88.5	14.6	1307	14	US-10-380-688-17
26	88.5	14.6	1388	15	US-10-093-463-138
27	88.5	14.6	1388	15	US-10-093-463-140
28	87	14.3	184	12	US-10-087-192-140
29	87	14.3	184	12	US-10-087-192-984
30	87	14.3	184	12	US-10-380-703-10
31	87	14.3	184	14	US-10-008-063-2
32	87	14.3	184	14	US-10-152-363A-60
33	87	14.3	266	12	US-10-380-703-7
34	87	14.3	426	12	US-10-257-174-44
35	87	14.3	426	12	US-10-343-357-6
36	85	14.0	872	16	US-10-437-963-154909
37	83	13.7	390	16	US-10-437-963-193427
38	81	13.3	484	16	US-10-437-963-182231
39	79.5	12.7	377	14	US-10-156-761-8962
40	77.5	12.7	261	16	US-10-640-863-5
41	77	12.6	448	16	US-10-218-654-31
42	76.5	12.6	250	14	US-10-262-439-31
43	76.5	12.6	276	14	US-10-218-654-26
44	76.5	12.6	276	14	US-10-262-439-26
45	76.5	12.6	14	14	US-10-262-439-26

ALIGNMENTS

RESULT 1
US-10-062-831-59
Sequence 59, Application US/10062831
Publication No. US20030105297A1
GENERAL INFORMATION:
APPLICANT: Steven M. Ruben, et al.
TITLE OF INVENTION: 32 Human Secreted Proteins
FILE REFERENCE: P2006P1
CURRENT APPLICATION NUMBER: US/10/062,831
CURRENT FILING DATE: 2002-02-05
PRIOR APPLICATION NUMBER: 09/690,454
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: PCT/US98/10868
PRIOR FILING DATE: May 28, 1998
PRIOR APPLICATION NUMBER: 60/044,039
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/050,935
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/048,093
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/048,190
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/048,356
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/056,250
PRIOR FILING DATE: August 29, 1997
PRIOR APPLICATION NUMBER: 60/056,296
PRIOR FILING DATE: August 29, 1997
PRIOR APPLICATION NUMBER: 60/056,293
PRIOR FILING DATE: August 29, 1997
NUMBER OF SEQ ID NOS: 229
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 59
LENGTH: 114
TYPE: PRT
ORGANISM: Homo sapiens

Sequence 139, App
Sequence 284, App
Sequence 228, App
Sequence 4, Appl
Sequence 7, Appl
Sequence 6, Appl
Sequence 14, Appl
Sequence 5, Appl
Sequence 17, Appl
Sequence 138, App
Sequence 140, App
Sequence 984, App
Sequence 10, Appl
Sequence 2, Appl
Sequence 7, Appl
Sequence 44, Appl
Sequence 6, Appl
Sequence 11595,
Sequence 193427,
Sequence 162231,
Sequence 8962, Ap
Sequence 181815,
Sequence 5, Appl
Sequence 31, Appl
Sequence 31, Appl
Sequence 26, Appl
Sequence 26, Appl

FEATURE:
NAME/KEY: SITE
LOCATION: (114)
OTHER INFORMATION: Xaa equals stop translation
US-10-062-831-59

Query Match 99.8%; Score 607; DB 14; Length 114;
Best Local Similarity 100.0%; Pred. No. 2,2e-51;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARGSLRLRLIVLIGLWALILRSVAGEAPGTAPCSRSGSSWSADLDKCMDCASCRRAPH 60
DB 1 MARGSLRLRLIVLIGLWALILRSVAGEAPGTAPCSRSGSSWSADLDKCMDCASCRRAPH 60
QY 61 SDFCLGCAAPAPAPFRLIWPILIGALSLTFVLGLSGFLVWRRCRRERSPP 113
DB 61 SDFCLGCAAPAPAPFRLIWPILIGALSLTFVLGLSGFLVWRRCRRERSPP 113

RESULT 2

US-10-062-599-59
Sequence 59, Application US/10062599
Publication No. US20030195346A1
GENERAL INFORMATION:
APPLICANT: Steven M. Ruben, et al.
TITLE OF INVENTION: 32 Human Secreted Proteins
FILE REFERENCE: P2006P1
CURRENT APPLICATION NUMBER: US/10/062,599
CURRENT FILING DATE: 2002-02-05
PRIOR APPLICATION NUMBER: 09/690,454
PRIOR FILING DATE: 2000-10-18
PRIOR APPLICATION NUMBER: 09/189,144
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: 60/044,039
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/048,093
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/048,190
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/050,935
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/048,101
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/048,356
PRIOR FILING DATE: May 30, 1997
PRIOR APPLICATION NUMBER: 60/056,250
PRIOR FILING DATE: August 29, 1997
PRIOR APPLICATION NUMBER: 60/056,296
PRIOR FILING DATE: August 29, 1997
PRIOR APPLICATION NUMBER: 60/056,293
PRIOR FILING DATE: August 29, 1997
NUMBER OF SEQ ID NOS: 229
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 59
LENGTH: 114
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (114)
OTHER INFORMATION: Xaa equals stop translation
US-10-062-599-59

Query Match 99.8%; Score 607; DB 14; Length 114;
Best Local Similarity 100.0%; Pred. No. 2,2e-51;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARGSLRLRLIVLIGLWALILRSVAGEAPGTAPCSRSGSSWSADLDKCMDCASCRRAPH 60
DB 1 MARGSLRLRLIVLIGLWALILRSVAGEAPGTAPCSRSGSSWSADLDKCMDCASCRRAPH 60
QY 61 SDFCLGCAAPAPAPFRLIWPILIGALSLTFVLGLSGFLVWRRCRRERSPP 113
DB 61 SDFCLGCAAPAPAPFRLIWPILIGALSLTFVLGLSGFLVWRRCRRERSPP 113

DB 61 SDFCLGCAAPAPAPFRLIWPILIGALSLTFVLGLSGFLVWRRCRRERSPP 113

RESULT 3

US-09-742-454A-4
Sequence 4, Application US/09742454A
Patent No. US20020041876A1
GENERAL INFORMATION:
APPLICANT: WILEY, Steven R.
TITLE OF INVENTION: TWEAK Receptor
FILE REFERENCE: 2968-B
CURRENT APPLICATION NUMBER: US/09/742,454A
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 60/172,878
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: 60/203,347
PRIOR FILING DATE: 2000-05-10
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 129
TYPE: PRT
ORGANISM: Homo sapiens
US-09-742-454A-4

Query Match 93.9%; Score 571; DB 9; Length 129;
Best Local Similarity 94.7%; Pred. No. 8e-48;
Matches 107; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MARGSLRLRLIVLIGLWALILRSVAGEAPGTAPCSRSGSSWSADLDKCMDCASCRRAPH 60
DB 1 MARGSLRLRLIVLIGLWALILRSVAGEAPGTAPCSRSGSSWSADLDKCMDCASCRRAPH 60
QY 61 SDFCLGCAAPAPAPFRLIWPILIGALSLTFVLGLSGFLVWRRCRRERSPP 113
DB 61 SDFCLGCAAPAPAPFRLIWPILIGALSLTFVLGLSGFLVWRRCRRERSPP 113

RESULT 4

US-09-883-777-4
Sequence 4, Application US/09883777
Patent No. US20020110853A1
GENERAL INFORMATION:
APPLICANT: WILEY, Steven R.
TITLE OF INVENTION: TWEAK RECEPTOR
FILE REFERENCE: 2968-C
CURRENT APPLICATION NUMBER: US/09/883,777
CURRENT FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: US 60/172,878
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: US 60/203,347
PRIOR FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: PCT/US00/34755
PRIOR FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 09/742,454
PRIOR FILING DATE: 2000-12-19
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 129
TYPE: PRT
ORGANISM: homo sapiens
US-09-883-777-4

Query Match 93.9%; Score 571; DB 9; Length 129;
Best Local Similarity 94.7%; Pred. No. 8e-48;
Matches 107; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MARGSLRLRLIVLIGLWALILRSVAGEAPGTAPCSRSGSSWSADLDKCMDCASCRRAPH 60
DB 1 MARGSLRLRLIVLIGLWALILRSVAGEAPGTAPCSRSGSSWSADLDKCMDCASCRRAPH 60
QY 61 SDFCLGCAAPAPAPFRLIWPILIGALSLTFVLGLSGFLVWRRCRRERSPP 113
DB 61 SDFCLGCAAPAPAPFRLIWPILIGALSLTFVLGLSGFLVWRRCRRERSPP 113

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2004, 10:00:13 ; Search time 18 Seconds
(without alignments)
326.964 Million cell updates/sec

Title: US-10-062-831-59

Perfect score: 608

Sequence: 1 MARGSLRLRLVIGLWLA.....LSGLVWRKCRERSPPPX 114

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfltest1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	607	99.8	114	4	US-09-690-454-59
2	274.5	45.1	112	4	US-09-489-847-139
3	274.5	45.1	155	4	US-09-489-847-284
4	274.5	45.1	156	4	US-09-489-847-228
5	96.5	15.9	248	4	US-09-252-991A-29249
6	88.5	14.6	400	4	US-09-252-991A-26145
7	87.5	14.4	631	4	US-09-252-991A-20063
8	83.5	13.7	152	4	US-09-252-991A-31619
9	80.5	13.2	249	4	US-09-252-991A-29850
10	76.5	12.6	250	4	US-09-322-409-31
11	76.5	12.6	250	4	US-09-451-527-31
12	76.5	12.6	276	4	US-09-322-409-26
13	76.5	12.6	276	4	US-09-451-527-26
14	75	12.3	334	4	US-09-252-991A-18795
15	74.5	12.3	305	4	US-09-252-991A-21147
16	73.5	12.1	187	4	US-09-199-637A-287
17	73.5	12.1	187	4	US-09-252-991A-21454
18	72	11.8	1278	4	US-09-462-136-2
19	72	11.8	3724	2	US-08-804-227C-10
20	72	11.8	3724	2	US-08-804-198-4
21	71.5	11.8	478	4	US-09-252-991A-22078
22	70.5	11.6	402	4	US-09-252-991A-18195
23	69.5	11.4	176	4	US-09-252-991A-25290
24	69	11.3	152	4	US-09-252-991A-24730
25	69	11.3	153	4	US-09-252-991A-20688
26	69	11.3	215	3	US-09-320-528-104
27	69	11.3	511	4	US-09-252-991A-28223

28	68.5	11.3	127	4	US-09-489-039A-10884	Sequence 10884, A
29	68.5	11.3	205	2	US-08-775-009-37	Sequence 37, Appl
30	68.5	11.3	429	1	US-07-964-589-2	Sequence 2, Appl1
31	68.5	11.3	429	5	PCT-US93-02024-2	Sequence 2, Appl1
32	68	11.2	139	4	US-09-252-991A-17414	Sequence 17414, A
33	68	11.2	144	4	US-09-252-991A-17313	Sequence 17313, A
34	68	11.2	204	4	US-09-252-991A-29996	Sequence 29996, A
35	68	11.2	282	4	US-09-252-991A-29124	Sequence 29124, A
36	68	11.2	366	4	US-09-252-991A-31958	Sequence 31958, A
37	68	11.2	775	4	US-09-252-991A-22300	Sequence 22300, A
38	67.5	11.1	174	3	US-09-383-586-12	Sequence 12, Appl
39	67.5	11.1	268	4	US-09-322-409-23	Sequence 23, Appl
40	67.5	11.1	268	4	US-09-451-527-23	Sequence 23, Appl
41	67.5	11.1	294	4	US-09-322-409-7	Sequence 7, Appl1
42	67.5	11.1	294	4	US-09-451-527-7	Sequence 7, Appl1
43	67.5	11.1	359	3	US-09-413-814-90	Sequence 90, Appl
44	67.5	11.1	361	3	US-09-413-814-77	Sequence 77, Appl
45	67	11.0	112	4	US-09-252-991A-21707	Sequence 21707, A

ALIGNMENTS

```
RESULT 1
US-09-690-454-59
; Sequence 59, Application US/09690454
; Patent No. 6531447
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: P2006P1
; CURRENT APPLICATION NUMBER: US/09/690,454
; CURRENT FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,150
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 59
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (114)
; OTHER INFORMATION: Xaa equals stop translation
US-09-690-454-59
```

```
Query Match 99.8% Score 607; DB 4; Length 114;
Best Local Similarity 100.0%; Pred. No. 3.5e-59;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MARGSLRLRLVIGLWLAIRSVAGCAPGTABCSRSSWSADLDKCMDCASGRARH 60
DB 1 MARGSLRLRLVIGLWLAIRSVAGCAPGTABCSRSSWSADLDKCMDCASGRARH 60
```

```
QY      61 SDFCLGCAAPPAFRLMLPILGASLTFVLGLISGFLVWRRCRRSSPP 113
Db      61 SDFCLGCAAPPAFRLMLPILGASLTFVLGLISGFLVWRRCRRSSPP 113
```

RESULT 2

```

US-09-489-847-139
Sequence 139, Application US/09489847
Patent No. 6476195
GENERAL INFORMATION:
APPLICANT: Rosen et al
TITLE OF INVENTION: 98 Human Secreted Proteins
FILE REFERENCE: P2051PI
CURRENT APPLICATION NUMBER: US/09/489, 847
CURRENT FILING DATE: 2000-01-24
EARLIER APPLICATION NUMBER: PCT/US99/17130
EARLIER FILING DATE: 1999-07-29
EARLIER APPLICATION NUMBER: 60/094, 657
EARLIER FILING DATE: 1998-07-30
EARLIER APPLICATION NUMBER: 60/095, 486
EARLIER FILING DATE: 1998-08-05
EARLIER APPLICATION NUMBER: 60/096, 319
EARLIER FILING DATE: 1998-08-12
EARLIER APPLICATION NUMBER: 60/095, 454
EARLIER FILING DATE: 1998-08-06
EARLIER APPLICATION NUMBER: 60/095, 455
EARLIER FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 376
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 139
LENGTH: 112
TYPE: prt
ORGANISM: Homo sapiens
US-09-489-847-139

```

Query Match	45.1%;	Score 274.5;	DB 4;	Length 112;
Best Local Similarity	96.4%;	Pred. No. 8.9e-23;		
Matches	54;	Conservative	1;	Mismatches 0;
				Indels 1;
				Gaps 1

Dy 1 MARGSLRLRLVLTGLWTLILRSVAGEQAPGTA PCSRSSWSADLDKCMDC A-SC 55
		:		

Dd 1 MARGSLRLRLVLTGLWTLILRSVAGEQAPGTA PCSRSSWSADLDKCMDC STSC 56

RESULT 3

```

US-09-489-847-284
Sequence 284, Application US/09489847
Patent No. 6476195
GENERAL INFORMATION:
APPLICANT: Rosen et al
TITLE OF INVENTION: 98 Human Secreted Proteins
FILE REFERENCE: P2031P
CURRENT APPLICATION NUMBER: US/09/489, 847
CURRENT FILING DATE: 2000-01-24
EARLIER APPLICATION NUMBER: PCT/US99/17130
EARLIER FILING DATE: 1999-07-29
EARLIER APPLICATION NUMBER: 60/094, 657
EARLIER FILING DATE: 1998-07-30
EARLIER APPLICATION NUMBER: 60/095, 486
EARLIER FILING DATE: 1998-08-05
EARLIER APPLICATION NUMBER: 60/096, 319
EARLIER FILING DATE: 1998-08-12
EARLIER APPLICATION NUMBER: 60/095, 454
EARLIER FILING DATE: 1998-08-06
EARLIER APPLICATION NUMBER: 60/095, 455
EARLIER FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 376
SOFTWARE: PatentIn Ver. 2. 0
SEQ ID NO 284
LENGTH: 155
TYPE: PRT
ORGANISM: Homo sapiens
US-09-489-847-284

```

Query Match	45.1%;	Score 274.5;	DB 4;	Length 155;
Best Local Similarity	-96.4%;	Pred. No. 1.3e-22;		
Matches 54;	Conservative 1;	Mismatches 0;	Indels 1;	Gaps 1;

[illegible]

TEST	RESULT	DATE
1	2	3
4	5	6
7	8	9
10	11	12
13	14	15
16	17	18
19	20	21
22	23	24
25	26	27
28	29	30
31	32	33
34	35	36
37	38	39
40	41	42
43	44	45
46	47	48
49	50	51
52	53	54
55	56	57
58	59	60
61	62	63
64	65	66
67	68	69
70	71	72
73	74	75
76	77	78
79	80	81
82	83	84
85	86	87
88	89	90
91	92	93
94	95	96
97	98	99
100	101	102
103	104	105
106	107	108
109	110	111
112	113	114
115	116	117
118	119	120
121	122	123
124	125	126
127	128	129
130	131	132
133	134	135
136	137	138
139	140	141
142	143	144
145	146	147
148	149	150
151	152	153
154	155	156
157	158	159
160	161	162
163	164	165
166	167	168
169	170	171
172	173	174
175	176	177
178	179	180
181	182	183
184	185	186
187	188	189
190	191	192
193	194	195
196	197	198
199	200	201
202	203	204
205	206	207
208	209	210
211	212	213
214	215	216
217	218	219
220	221	222
223	224	225
226	227	228
229	230	231
232	233	234
235	236	237
238	239	240
241	242	243
244	245	246
247	248	249
250	251	252
253	254	255
256	257	258
259	260	261
262	263	264
265	266	267
268	269	270
271	272	273
274	275	276
277	278	279
280	281	282
283	284	285
286	287	288
289	290	291
292	293	294
295	296	297
298	299	300
301	302	303
304	305	306
307	308	309
310	311	312
313	314	315
316	317	318
319	320	321
322	323	324
325	326	327
328	329	330
331	332	333
334	335	336
337	338	339
340	341	342
343	344	345
346	347	348
349	350	351
352	353	354
355	356	357
358	359	360
361	362	363
364	365	366
367		

```

US-09-489-847-228
? Sequence 228, Application US/09489847
? Patent No. 6476195
? GENERAL INFORMATION:
? APPLICANT: Rosen et al
? TITLE OF INVENTION: 98 Human Secreted Proteins
? FILE REFERENCE: P2031p1
? CURRENT APPLICATION NUMBER: US/09/489,847
? CURRENT FILING DATE: 2000-01-24
? EARLIER APPLICATION NUMBER: PCT/US99/17130
? EARLIER FILING DATE: 1999-07-29
? EARLIER APPLICATION NUMBER: 60/094,657
? EARLIER FILING DATE: 1998-07-30
? EARLIER APPLICATION NUMBER: 60/095,486
? EARLIER FILING DATE: 1998-08-05
? EARLIER APPLICATION NUMBER: 60/096,319
? EARLIER FILING DATE: 1998-08-12
? EARLIER APPLICATION NUMBER: 60/095,454
? EARLIER FILING DATE: 1998-08-06
? EARLIER APPLICATION NUMBER: 60/095,455
? EARLIER FILING DATE: 1998-08-06
? NUMBER OF SEQ ID NOS: 376
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 228
? LENGTH: 156
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURES:
? NAME/KEY: SITE
? LOCATION: (156)
? OTHER INFORMATION: Xaa equals stop translation
? US-09-489-847-228

```

Query Match	45.1%;	Score 274.5;	DB 4;	Length 156;
Best Local Similarity	96.4%;	Pred. No. 1.3e-22;		
Matches 54;	Conservative 1;	Mismatches 0;	Indels 1;	Gaps 1;

[illegible]

RESULT 5

```

US-09-252-991A29249
/ Sequence 29249, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 29249
/ LENGTH: 248
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2004, 09:55:12 ; Search time 54 Seconds

(without alignments)
596,489 Million cell updates/sec

Title: US-10-062-831-59
Perfect score: 608

Sequence: 1 MARGSLRLRLILVGLHMLA.....LSGFLVWRRCRRRSPPEX 114

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003ds:.*
8: geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	607	99.8	114	2	AAW73409 Human sec
2	571	93.9	129	2	AAW88506 Human liv
3	571	93.9	129	3	AAW57940 Human tra
4	571	93.9	129	4	AAU03498 Human TWE
5	571	93.9	129	5	ABP61512 Human NF
6	571	93.9	129	5	AAU79827 Human TYP
7	571	93.9	129	6	ADA56889 Human sec
8	571	93.9	129	6	ABU56716 Lung can
9	571	93.9	129	7	ADG74112 Human sec
10	571	93.9	129	7	ADD37867 Human sec
11	571	93.9	129	7	ADD89033 Human sec
12	450	74.0	129	5	ADD89033 TAT274. 1
13	433	71.2	309	4	AAU03500 Human TWE
14	274.5	45.1	112	6	AAW91463 Human TWE
15	274.5	45.1	112	6	ADA57390 Human sec
16	274.5	45.1	112	7	ADG74462 Human sec
17	274.5	45.1	112	7	ADD38025 Human sec
18	274.5	45.1	112	7	ADD38025 Human sec
19	274.5	45.1	112	7	ADD38025 Human sec
20	274.5	45.1	112	7	ADD38025 Human sec
21	274.5	45.1	112	7	ADD38025 Human sec
22	274.5	45.1	112	7	ADD38025 Human sec
23	191	31.4	156	3	AAW91552 Human sec
24	97.5	16.0	32	6	ADA49370 Human Fnl
25	97.5	16.0	171	7	ADB90668 TALL-IR g
			171	7	ADB90665 TALL-IR g

26	97.5	16.0	185	7	ADB90663
27	97	16.0	170	7	ADB90667
28	96.5	15.9	166	7	ADB90675
29	94.5	15.5	185	5	AAE22269
30	92.5	15.2	185	5	AAE22267
31	92.5	15.2	185	5	AAE22266
32	91.5	15.0	185	5	AAE22270
33	91.5	15.0	185	5	AAE22271
34	89.5	14.7	185	5	AAE22268
35	88.5	14.6	185	5	AAE22242
36	88.5	14.6	1307	7	AAE39974
37	88.5	14.6	1388	5	ABP70118
38	88.5	14.6	1388	5	ABP70117
39	87.5	14.4	185	7	ABR82287
40	87	14.3	184	5	ABR81483
41	87	14.3	184	6	AAE35227
42	87	14.3	184	6	ADA49363
43	87	14.3	184	6	ABP97721
44	87	14.3	184	6	ABR61767
45	87	14.3	266	5	AAE22243

ALIGNMENTS

RESULT 1	AAW73409	standard; protein; 114 AA.
ID	AAW73409	
XX	AAW73409;	
DT	19-FEB-1999	(first entry)
XX		
DE	Human secreted protein encoded by Gene No. 13.	
XX		
KW	Secreted protein; human; protein therapy; gene therapy; blood disorder;	
KW	pathological condition; diagnosis; cancer; neurological disorder;	
KW	developmental abnormality; foetal deficiency; leukaemia; hepatic disease;	
KW	immune system disorder; Alzheimer's disease; cognitive disorder;	
KW	schizophrenia; prostate disease; autoimmune disorder; AIDS.	
XX		
OS	Homo sapiens.	
XX		
EH	Key	Location/Qualifiers
FT	Misc-difference 114	/note="unspecified amino acid"
XX		
XX	WO9854206-A1.	
PD	03-DEC-1998.	
XX		
XX	28-MAY-1998;	98WO-US010868.
XX	30-MAY-1997;	97US-0044039P.
XX	30-MAY-1997;	97US-0048093P.
XX	30-MAY-1997;	97US-0048101P.
XX	30-MAY-1997;	97US-0048190P.
XX	30-MAY-1997;	97US-0048356P.
XX	30-MAY-1997;	97US-0050935P.
XX	29-AUG-1997;	97US-0056250P.
XX	29-AUG-1997;	97US-0056295P.
XX	29-AUG-1997;	97US-0056296P.
XX		
XX	(HUMA-) HUMAN GENOME SCI INC.	
XX		
XX	Ruben SM, Rosen CA, Carter KC, Dillon PJ, Endress GA, Yu G;	
XX	Ni U, Feng P;	
XX	WP.; 1999-070209/06.	
XX	N-PSDB; AAW08823.	
XX		
XX	New isolated human genes - useful for diagnosis and treatment of, e.g.	
XX	cancers, neurological disorders, immune diseases, developmental disorders	

PT or blood disorders.

XX
PS Claim 11; Page 153; 189pp; English.

CC This sequence is encoded by a cDNA of the invention, designated Gene No.
CC 13. This sequence represents a human secreted protein, and is expressed
CC in keratinocytes and to a lesser extent in endothelial cells and
CC placenta. The DNA sequences of the invention and their corresponding
CC secreted polypeptides are useful for preventing, treating or ameliorating
CC medical conditions, e.g. by protein or gene therapy. Also pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the DNA sequences. Specific uses are described for each of the DNA
CC sequences and the encoded proteins, based on which tissues they are most
CC highly expressed in, and include developing products for the diagnosis or
CC treatment of cancer, tumors, neurological disorders, developmental
CC abnormalities and foetal deficiencies, blood disorders, leukaemias,
CC diseases of the immune system (including allergies or asthma), hepatic
CC diseases, Alzheimer's and cognitive disorders, schizophrenia, prostate
CC diseases, autoimmune disorders and AIDS. The polypeptides are also useful
CC for identifying their binding partners

SO Sequence 114 AA;

Query Match 99.8%; Score 607; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 1.3e-55;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARGSLRLRLRLVLTGLMTALRSVAGEQAPGTAPCSRGSWSADLDKMDCASCRARPH 60
DB 1 MARGSLRLRLRLVLTGLMTALRSVAGEQAPGTAPCSRGSWSADLDKMDCASCRARPH 60

QY 61 SDFCLGCAAPAPAPFRLMLPILIGALSLTFVLGLSGFLVWRRCRRERSPP 113
DB 61 SDFCLGCAAPAPAPFRLMLPILIGALSLTFVLGLSGFLVWRRCRRERSPP 113

RESULT 2
AAW88506
ID AAW88506 standard; protein; 129 AA.

XX
AC AAW88506;
XX
DT 30-MAR-1999 (first entry)

XX Human liver clone HP10432-encoded membrane protein.

XX Transmembrane protein; HP10432; human; liver.

XX Homo sapiens.

XX MO9855508-A2.

XX 10-DEC-1998.

XX 03-JUN-1998; 98WO-JP002445.

XX 03-JUN-1997; 97JP-00144948.

XX (SAGA) SAGAMI CHEM RES CENTRE.
PA (PROT-) PROTEGENE INC.

XX Kato S, Sekine S, Yamaguchi T;

XX WPI; 1999-045730/04.

XX N-PSDB; AAV84374.

PT New human proteins containing transmembrane domains and their encoding
PT sequences - useful in the preparation of antibodies and large-scale
PT protein production, gene diagnosis, and gene therapy.

PS Claim 1; Page 152-153; 178pp; English.

CC This is the amino acid sequence of a transmembrane protein encoded by
CC human liver cDNA clone HP10432 (see AAW84374). The encoded protein has a
CC signal-like N-terminal region and one internal transmembrane domain. The
CC invention provides nucleotide sequences (see AAW84359-76) coding for 18
CC polynucleotides, and eukaryotic cells containing the vectors. The
CC proteins can be used as antigens or as compositions in the preparation of
CC antibodies against the proteins. The polynucleotides can be used as
CC probes for gene diagnosis, and as gene sources for gene therapy and large
CC scale production of proteins encoded by the cDNA. The host cells are
CC used for the detection of ligands corresponding to the expressed
CC proteins, and the screening of low mol.wt. medicines

SO Sequence 129 AA;

Query Match 93.9%; Score 571; DB 2; Length 129;
Best Local Similarity 94.7%; Pred. No. 8.4e-52;
Matches 107; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MARGSLRLRLRLVLTGLMTALRSVAGEQAPGTAPCSRGSWSADLDKMDCASCRARPH 60
DB 1 MARGSLRLRLRLVLTGLMTALRSVAGEQAPGTAPCSRGSWSADLDKMDCASCRARPH 60

QY 61 SDFCLGCAAPAPAPFRLMLPILIGALSLTFVLGLSGFLVWRRCRRERSPP 113
DB 61 SDFCLGCAAPAPAPFRLMLPILIGALSLTFVLGLSGFLVWRRCRRERSPP 113

RESULT 3
AAV57940
ID AAV57940 standard; protein; 129 AA.

XX AAV57940;

XX 23-MAR-2000 (first entry)

XX Human transmembrane protein HTMPN-64.

XX Human; transmembrane protein; HTMPN; diagnosis; immunospecific;

XX antiproliferative; neuroprotective; immune disorder;

XX reproductive disorder; smooth muscle disorder; neurological disorder;

XX gastrointestinal disorder; developmental disorder;

XX cell proliferative disorder.

XX Homo sapiens.

XX MO9961471-A2.

XX 02-DEC-1999.

XX 28-MAY-1999; 99WO-US011904.

XX 29-MAY-1998; 98US-0087260P.

XX 02-JUL-1998; 98US-0091674P.

XX 02-OCT-1998; 98US-0102954P.

XX 24-NOV-1998; 98US-0109869P.

XX (INCYTE) INCYTE PHARM INC.

XX Tang YT, Lal P, Hillman JL, Yue H, Guegler KT, Corley NC;

XX Bandman O, Patterson C, Gorgone GA, Kaser MK, Baughn MR, Au-Young J;

XX WPI; 2000-072605/06.

XX N-PSDB; AAZ56761.

PT Proteins, polynucleotides, vectors, host cells and antibodies used to
PT diagnose, treat or prevent immune, reproductive, smooth muscle,
PT neurological, gastrointestinal, developmental and cell proliferative
PT disorders.

PS Claim 1; Page 163; 229pp; English.

PS AAZ56698 to AAZ56776 encode AAV57877 to AAV57955 which represent human

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2004, 09:55:37 ; Search time 13 Seconds

(without alignments)
456.615 Million cell updates/sec

Title: US-10-062-831-59
Perfect score: 608

Sequence: 1 MARGSLRRLRLVLGLMLA.....LSGLVWRRCRRRSPPX 114

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	571	93.9	129	1	Q9np84 homo sapien
2	445	73.2	129	1	Q9c75 m tumor nec
3	87	14.3	184	1	Q96vj3 homo sapien
4	78	12.8	448	1	Q9wv18 ratu
5	77	12.7	448	1	Q9wv18 ratu
6	76.5	12.6	245	1	Q9wv18 ratu
7	74	12.2	548	1	Q9wv18 ratu
8	73.5	12.1	81	1	Q9wv18 ratu
9	72	11.8	1278	1	Q9wv18 ratu
10	72	11.8	2871	1	Q9wv18 ratu
11	70	11.5	301	1	Q9wv18 ratu
12	70	11.5	443	1	Q9wv18 ratu
13	70	11.5	558	1	Q9wv18 ratu
14	69.5	11.4	495	1	Q9wv18 ratu
15	69.5	11.3	180	1	Q9wv18 ratu
16	69	11.3	314	1	Q9wv18 ratu
17	69	11.3	324	1	Q9wv18 ratu
18	68.5	11.3	205	1	Q9wv18 ratu
19	68.5	11.3	431	1	Q9wv18 ratu
20	67.5	11.1	368	1	Q9wv18 ratu
21	67.5	11.1	1008	1	Q9wv18 ratu
22	67.5	11.1	1013	1	Q9wv18 ratu
23	67	11.0	279	1	Q9wv18 ratu
24	67	11.0	319	1	Q9wv18 ratu
25	67	11.0	386	1	Q9wv18 ratu
26	67	11.0	1005	1	Q9wv18 ratu
27	67	11.0	2871	1	Q9wv18 ratu
28	66.5	10.9	443	1	Q9wv18 ratu
29	66	10.9	394	1	Q9wv18 ratu
30	66	10.9	992	1	Q9wv18 ratu
31	66	10.9	1238	1	Q9wv18 ratu
32	66	10.9	2318	1	Q9wv18 ratu
33	66	10.9	2319	1	Q9wv18 ratu

34	65.5	10.8	245	1	FLIP ECOLI	P33133 escherichia
35	65.5	10.8	2321	1	NTC3 HUMAN	Q9um47 homo sapien
36	65	10.7	453	1	HRA3 HUMAN	P83110 homo sapien
37	65	10.7	1014	1	EPB6 MOUSE	O08644 mus musculu
38	65	10.7	1445	1	PTP6 HUMAN	P23470 homo sapien
39	64.5	10.6	112	1	PLA6 MOUSE	Q9148 mus musculu
40	64.5	10.6	123	1	VST1 HEVME	O03499 hepatitis e
41	64.5	10.6	300	1	TR6B HUMAN	O05407 homo sapien
42	64.5	10.6	443	1	FLB4 HUMAN	O05967 homo sapien
43	64	10.5	115	1	PLA6 HUMAN	Q9nzt1 homo sapien
44	64	10.5	1210	1	EGFR MOUSE	O01279 mus musculu
45	64	10.5	1227	1	B3A3 MOUSE	P16283 mus musculu

ALIGNMENTS

RESULT 1
ID FN14_HUMAN STANDARD; PRT; 129 AA.
AC Q9np84; Q9HCS0;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Tumor necrosis factor receptor superfamily member Fn14 precursor
DE (Fibroblast growth factor-inducible immediate-early response protein
DE 14) (FGF-inducible 14) (Tweak-receptor) (TweakR).
GN TNFRSF12A OR FN14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Placenta;
RX MEDLINE=20216634; PubMed=10751351;
RA Feng S.-L.Y., Guo Y., Factor V.M., Thorgeirsson S.S., Bell D.W.,
RA Testa F.R., Peifley K.A., Winkles J.A.;
RT "The Fn14 immediate-early response gene is induced during liver
RT regeneration and highly expressed in both human and murine
RT hepatocellular carcinoma";
RL Am. J. Pathol. 156:1253-1261 (2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC Tanaka S., Sugimachi K.;
RX "Human homologue of Fn14";
RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=uterus;
RX MEDLINE=22386257; PubMed=12477932;
RA Strassberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang L.J., Hsieh F.,
RA Diatchenko L., Martins K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Scherker A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903 (2002).
RN [4]
RP FUNCTION.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2004, 09:58:42 ; Search time 39 Seconds

(without alignments)
922.284 Million cell updates/sec

Title: US-10-062-831-59

Sequence: 1 MARGSLRLRLRLVGLWMLA.....LSGFLVWRRCRRRSPPPX 114

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	453	74.5	129	11 Q80XX9	Q80XX9 ratius norv
2	87	14.3	333	4 Q9BZG3	Q9BZG3 homo sapien
3	87	14.3	426	4 Q9BZG2	Q9BZG2 homo sapien
4	83.5	13.7	610	10 Q943G8	Q943G8 oryza sativ
5	79.5	13.1	377	16 Q82NR0	Q82NR0 streptomyce
6	79	13.0	341	6 Q9N0R3	Q9N0R3 macaca fasc
7	78.5	12.9	332	10 Q42839	Q42839 hordeum vul
8	77.5	12.7	387	11 Q8VD70	Q8VD70 mus musculu
9	77.5	12.7	219	11 Q8BM15	Q8BM15 mus musculu
10	76.5	12.6	245	16 Q8ZSR3	Q8ZSR3 mus musculu
11	76.5	12.6	245	16 Q8ZSR3	Q8ZSR3 mus musculu
12	75.5	12.4	222	12 Q9QSE6	Q9QSE6 indian citr
13	75.5	12.4	436	16 Q7U2G8	Q7U2G8 mycobacteri
14	75.5	12.4	449	16 Q53668	Q53668 mycobacteri
15	75.5	12.4	449	16 Q53668	Q53668 mycobacteri
16	75.5	12.4	635	5 Q18288	Q18288 caenorhabdi

17	75	12.3	370	4 Q96KN9	Q96KN9 homo sapien
18	75	12.3	411	16 Q8CXG1	Q8CXG1 oceanobacil
19	74.5	12.3	202	12 Q80MP4	Q80MP4 indian citr
20	74.5	12.3	330	7 Q860W5	Q860W5 mus musculu
21	74.5	12.3	519	16 Q9A9X4	Q9A9X4 caulobacter
22	74	12.2	768	5 P90890	P90890 caenorhabdi
23	73.5	12.1	409	16 Q82FA3	Q82FA3 streptomyce
24	73.5	12.1	730	4 Q9P2F7	Q9P2F7 homo sapien
25	73	12.0	237	10 Q04393	Q04393 hordeum vul
26	73	12.0	243	16 Q8BQD6	Q8BQD6
27	73	12.0	370	4 Q8N2R7	Q8N2R7
28	73	12.0	768	13 Q96TH8	Q96TH8 cyprinus ca
29	72.5	11.9	206	4 Q96QAO	Q96QAO
30	72.5	11.9	242	16 Q8NNCO	Q8NNCO
31	72.5	11.9	329	4 Q9NND2	Q9NND2
32	72.5	11.9	346	16 Q9RSC0	Q9RSC0
33	72.5	11.9	365	4 Q9Y288	Q9Y288
34	72.5	11.9	370	4 Q9B2M8	Q9B2M8
35	72.5	11.9	445	2 Q7MX79	Q7MX79
36	72	11.8	368	13 Q9IAR7	Q9IAR7
37	72	11.8	721	5 Q81BV6	Q81BV6
38	71.5	11.8	210	4 Q95054	Q95054
39	71.5	11.8	379	10 Q8W393	Q8W393
40	71.5	11.8	2873	12 Q93072	Q93072
41	71	11.7	116	12 Q90631	Q90631
42	71	11.7	175	16 Q9S255	Q9S255
43	71	11.7	308	16 Q7MR57	Q7MR57
44	71	11.7	308	16 Q7W290	Q7W290
45	71	11.7	308	16 Q7VT68	Q7VT68

ALIGNMENTS

RESULT 1	Q80XX9	PRELIMINARY;	PRT;	129 AA.
ID	Q80XX9			
AC	Q80XX9;			
DT	01-JUN-2003 (TREMBLrel. 24, Created)			
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	Type 1 transmembrane protein FN14.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Mueller A.M., Giegerich G.;			
RT	"FN14, TNFRSF12A."			
RL	Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AY25102; AAP06753.1, ..			
DR	GO: GO:0016021; C: integral to membrane; IEA.			
KW	Transmembrane.			
SQ	SEQUENCE 129 AA; 13616 MW; 57143F979CD7AC77 CRC64;			

Query Match	74.5%	Score 453;	DB 11;	Length 129;
Best Local Similarity	76.1%;	Pred. NO. 1.9e-39;		
Matches	86;	Conservative	5;	Mismatches 22;
				Indels 0;
				Gaps 0;

QY	1	MARGSLRLRLRLVGLWMLA	IRSVAGQAPGTA	PCSGSSMSADIDKMCACGRAPH	60
DB	1	MARGSLRLRLRLVGLWMLA	IRSVAGQAPGTA	PCSGSSMSADIDKMCACGRAPH	60
QY	61	SDFLGCAAPPAFFRLIMPLIGALSITFYVGLSGFLVWRRCRRRSPPP			113
DB	61	SDFLGCAAPPAFFRLIMPLIGALSITFYVGLSGFLVWRRCRRRSPPP			113

RESULT 2
Q9BZG3
ID Q9BZG3
AC Q9BZG3;
PRELIMINARY; PRT; 333 AA.

```

DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Acid phosphatase variant 3.
GN ACPT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21309073; PubMed=11414767;
RA Yousef G.M., Diamandis M., Jung K., Diamandis E.P.;
RT "Molecular cloning of a novel human acid phosphatase gene (ACPT) that
RT is highly expressed in the testis.";
RL Genomics 74:385-395(2001).
DR EMBL; AF321918; AAK09393.1; -.
DR HSP; P15309; 2HPA.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR000560; HisAc_phsphtse.
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT 1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT 2; 1.
SQ SEQUENCE 333 AA; 35402 MW; 0C0237096567B30F CRC64;

Query Match 14.3%; Score 87; DB 4; Length 333;
Best Local Similarity 35.3%; Pred. No. 0.49;
Matches 30; Conservative 6; Mismatches 35; Indels 14; Gaps 5;

QY 31 PG-TAPCRSSMSADLDKMDCASCRARPHSDFLGCG--AAAPAPFRLMLPILGALS 87
DB 258 PGCPAPCPGGRFYQL-----TAPAPPAHGVSCHPYEAALPPAP--VPELLAGAVA 307
QY 88 IFEVLGLSGFLVWR-RCRERSSP 111
DB 308 VVALSLGILGILAMRGCLRALGSP 332

RESULT 3
Q9BZG2 PRELIMINARY; PRT; 426 AA.
ID Q9BZG2;
AC Q9BZG2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Acid phosphatase.
GN ACPT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21309073; PubMed=11414767;
RA Yousef G.M., Diamandis M., Jung K., Diamandis E.P.;
RT "Molecular cloning of a novel human acid phosphatase gene (ACPT) that
RT is highly expressed in the testis.";
RL Genomics 74:385-395(2001).
DR EMBL; AF321918; AAK09393.1; -.
DR HSP; P15309; 2HPA.
DR GeneW; HGNC:14376; ACPT.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR000560; HisAc_phsphtse.
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT 1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT 2; 1.
SQ SEQUENCE 426 AA; 46089 MW; BE930398041B061 CRC64;

Query Match 14.3%; Score 87; DB 4; Length 426;
Best Local Similarity 35.3%; Pred. No. 0.62;
Matches 30; Conservative 6; Mismatches 35; Indels 14; Gaps 5;

QY 31 PG-TAPCRSSMSADLDKMDCASCRARPHSDFLGCG--AAAPAPFRLMLPILGALS 87

```

```

DB 351 PGCPAPCPGGRFYQL-----TAPAPPAHGVSCHPYEAALPPAP--VPELLAGAVA 400
QY 88 IFEVLGLSGFLVWR-RCRERSSP 111
DB 401 VVALSLGILGILAMRGCLRALGSP 425

RESULT 4
Q943G8 PRELIMINARY; PRT; 610 AA.
ID Q943G8;
AC Q943G8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative receptor protein kinase.
RN P0046E05.12.
GN Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhacridae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone: P0046E05.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF003237; BAB67905.1; -.
DR Gramene; Q943G8; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004672; F:receptor activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; pkinase; 1.
DR PRODOM; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ATP; 1.
KW ATP-binding; Kinase; Receptor; Serine/threonine-protein kinase;
KW Transferase
SQ SEQUENCE 610 AA; 64405 MW; 2A7B3EBE83FA9303 CRC64;

Query Match 13.7%; Score 83.5; DB 10; Length 610;
Best Local Similarity 24.6%; Pred. No. 2;
Matches 34; Conservative 12; Mismatches 35; Indels 57; Gaps 4;

QY 30 APGTAPCRSSMSADLDKMD-----CASCRARPHSDP 63
DB 84 SPGSSSYTGSSSTAVTASTDADPNAGDAVPTSGADATPSSACRKAARCPKRPATD 142
QY 64 CLGCAAPAPFRLMLP-----ILGALSLTEVLGLSG 97
DB 143 -NVPASPPPRISLSPPTPTGSGASSGSKSNNGTVAVGVAAVVVGLAAGL 200
QY 98 --FLWRCRCRRSSPPP 113
DB 201 IYFVSKRRRRRQHPPP 218

RESULT 5
Q82N80 PRELIMINARY; PRT; 377 AA.
ID Q82N80;
AC Q82N80;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.

```


C;Function:
A;Description: may be involved in flagellar assembly; may be involved in export of flag
C;Superfamily: flagellar biosynthetic protein fljP
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-245/Product: probable export protein fljP #status predicted <MAT>
F;45-61/Domain: transmembrane #status predicted <TM1>
F;89-105/Domain: transmembrane #status predicted <TM2>
F;189-205/Domain: transmembrane #status predicted <TM3>
F;212-228/Domain: transmembrane #status predicted <TM4>

Query Match 12.6%; Score 76.5; DB 2; Length 245;
Best Local Similarity 30.8%; Pred. No. 4.3;
Matches 33; Conservative 9; Mismatches 38; Indels 27; Gaps 6;

OY 6 LRRLLRLVLGLMALLRSVAGEQAPG--TAP-CSSRSSMSADIDKCMDCASCARP--- 59
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MRRLFLSLAGLW--LSPPAALQLPGILISPLLAGGGQSMLSVQLLVFTLSLTLPAIL 58
| | | | | | | | | | | | | | | | | | | | | | | | | |

OY 60 --HSDP-----CLGCAAPPAPFRLLWPILLGALSTLPVL 92
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 59 LMMTSETRIIIVFGLLRNALGTSPAPPNQV----LIGLALFLTFPI 100
| | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 3
AD0753
Flagellar biosynthetic protein FljP [imported] - Salmonella enterica subsp. enterica sero
C;Species: Salmonella enterica subsp. enterica serovar Typh
A;Note: This species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AD0753
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
ch, T.; Connorton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AD0753
A>Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-245 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD0527.1; PID:G16503220; GSPDB:GN00176
C;Genetics:
A;Gene: STY2187
C;Superfamily: flagellar biosynthetic protein fljP

Query Match 12.6%; Score 76.5; DB 2; Length 245;
Best Local Similarity 30.8%; Pred. No. 4.3;
Matches 33; Conservative 9; Mismatches 38; Indels 27; Gaps 6;

OY 6 LRRLLRLVLGLMALLRSVAGEQAPG--TAP-CSSRSSMSADIDKCMDCASCARP--- 59
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MRRLFLSLAGLW--LSPPAALQLPGILISPLLAGGGQSMLSVQLLVFTLSLTLPAIL 58
| | | | | | | | | | | | | | | | | | | | | | | | | |

OY 60 --HSDP-----CLGCAAPPAPFRLLWPILLGALSTLPVL 92
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 59 LMMTSETRIIIVFGLLRNALGTSPAPPNQV----LIGLALFLTFPI 100
| | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
B70939
hypothetical protein RV0246 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: B70939
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Radnareem, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulistio, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: AV0500; MUID:98295987; PMID:9634230

```

A/Accession: B70939
A/Status: Preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-436 <COL>
A/Cross-references: GB:AL021929; GB:AL123456; NID:93242291; PIDN:CAAL7338.1; PID:el25244
C/Genetics:
A/Experimental source: strain H37RV
A/Gene: RV0246

Query Match      12.4%; Score 75.5; DB 2; Length 436;
Best Local Similarity 23.0%; Pred. No. 8.4;
Matches 35; Conservative 16; Mismatches 52; Indels 49; Gaps 5;

OY 1 MARGSLRRLRLILVGLMALILRSVAGEQA--PETAFCSSRGSS--WSADLDKMCDCASCR 56
DB 142 MPSTAAARILILTFVGGALITAVVATLSFVPDQHLSTNINHLMTAAVAMASALICR 201
OY 57 ARPHSDFCLGCAAPPAFFRLW-----PIIGG----- 84
DB 202 ALPFR--IVPRVHAAPGLHKLVYGVMTAIRTNGWRRRYLLVQVLGVSIVLSSPFSIRVA 259
OY 85 -----ALSLTFVLGLSGFLWRCR 105
DB 260 AVPGDQPEVVAVVLFEVCGILGIALMNRVR 291

RESULT 5
T20910
hypothetical protein ZK1010.9 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C/Accession: T20910; T27646
R/McMurray, A.
submitted to the EMBL Data Library, November 1996
A/Reference number: 219345
A/Accession: T20910
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-635 <W11>
A/Cross-references: EMBL:Z81503; PIDN:CA804115.1; GSPDB:GN00021; CESP:ZK1010.9
A/Experimental source: clone F14F7
R/Gardner, A.
submitted to the EMBL Data Library, November 1996
A/Reference number: Z20398
A/Accession: T27646
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-635 <W12>
A/Cross-references: EMBL:Z82083; PIDN:CA804975.1; GSPDB:GN00021; CESP:ZK1010.9
A/Experimental source: clone ZK1010
C/Genetics:
A/Gene: CESP:ZK1010.9
A/Map position: 3
A/Introns: 71/3; 184/2; 316/3; 403/1; 444/3; 547/3; 577/3
C/Superfamily: gamma-aminobutyric acid transporter

Query Match      12.4%; Score 75.5; DB 2; Length 635;
Best Local Similarity 36.6%; Pred. No. 11;
Matches 26; Conservative 10; Mismatches 28; Indels 7; Gaps 3;

OY 31 PGTAFCSGSSWSADLDKMCDCASGRAR--HSPFCLG--CAAPPAFFRLIMPIIGAL 86
DB 92 PTYAKNGSLF---LIVVVCGLIFAPPAIHMERFALGQYAAKSPPAFRMMPILEGV 148
OY 87 SLTFVLGLSG 97
DB 149 WMTCLVGAIIIG 159

RESULT 6
B87353
hypothetical protein CC0837 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus

```